

PRO polypeptide; mammal; tumour; cancer; human; cattle; horse; sheep;
dog; cat; pig; goat; rabbit; tumour necrosis factor alpha; TNF-alpha;
blood; chondrocyte cell; cell proliferation; cell differentiation; colon;
adrenal; lung; breast; prostate; rectum; cervix; liver; genetic disorder.

PA (GETH) GENENTECH INC.
XX Batou DL, Filvaroff E, Gerritsen ME, Goddard A, Godowski PV;
PI Grimaldi CJ, Gurney AL, Watanabe CK, Wood WI;
XX
DR WPI; 2001-183260/18.
DR N-PSDB; AAF92123.
XX
XX Eighty four nucleic acids encoding PRO polypeptides, useful in
PT molecular biology, including use as hybridization probes, and in
PT chromosome and gene mapping.
XX
XX Claim 12; Fig 132; 278pp; English.
XX
XX The present sequence is a human PRO polypeptide (secreted and
CC transmembrane). The PRO protein, and PRO agonists, PRO antagonists or
CC anti-PRO antibodies are useful for preparation of a medicament useful in
CC the treatment of a condition which is responsive to the PRO protein,
CC agonists, antagonists or anti-PRO antibodies. The PRO protein may also be
CC employed as molecular weight markers for protein electrophoresis. The PRO
CC coding sequence has applications in molecular biology, including use as
CC hybridisation probes, and in chromosome and gene mapping.
XX
XX Sequence 649 AA;
SQ
Query Match 100.0%; Score 3400; DB 22; Length 649;
Best Local Similarity 100.0%; Pred. No. 7.5e-290;
Matches 649; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MISAWSFLIGKIGLQVAPLSVMAKSCPSVCRDAGFIYCNDRFTSTPTGIPEDA 60
Db 1 MISAWSFLIGKIGLQVAPLSVMAKSCPSVCRDAGFIYCNDRFTSTPTGIPEDA 60
QY 61 TTYLQNNQINNAGIPSDKLNLLKVERIYLYHNSIDFPTNPKYKELHLOENIRIT 120
Db 61 TTYLQNNQINNAGIPSDKLNLLKVERIYLYHNSIDFPTNPKYKELHLOENIRIT 120
QY 121 YDSLSKIPYLEELHLDNDSVSAVSEIEGAFRDSNLYRLFLSRNHLSTIPWGLPTIEEL 180
Db 121 YDSLSKIPYLEELHLDNDSVSAVSEIEGAFRDSNLYRLFLSRNHLSTIPWGLPTIEEL 180
QY 181 RDDNRISTSSPSLOGLSLKRLVLDGNLLNHHGLDKFFNLVNLSTELSVRSLTAA 240
Db 181 RDDNRISTSSPSLOGLSLKRLVLDGNLLNHHGLDKFFNLVNLSTELSVRSLTAA 240
QY 241 PVNLPFTNLRKLYLQDNHINRVPNPAFSLYRLQYRLDMSNNLSNLPQGFDDLDNITQL 300
Db 241 PVNLPFTNLRKLYLQDNHINRVPNPAFSLYRLQYRLDMSNNLSNLPQGFDDLDNITQL 300
QY 301 ILRNNPWCCKMKWVRDMLQSLPVKVNVRGLMCOAPEKVRGMAIKDLNAELFDCKDSGI 360
Db 301 ILRNNPWCCKMKWVRDMLQSLPVKVNVRGLMCOAPEKVRGMAIKDLNAELFDCKDSGI 360
QY 361 VSTQITTAIPNVYPAQGWAPVTKPDINKPKLTKDOQTGSPSRKTIITIVKSVTS 420
Db 361 VSTQITTAIPNVYPAQGWAPVTKPDINKPKLTKDOQTGSPSRKTIITIVKSVTS 420
QY 421 DTIHSWKALPMTALFSLKSLGHSPAFGSITETIYTGSELYLTALPDSPYKVCWV 480
Db 421 DTIHSWKALPMTALFSLKSLGHSPAFGSITETIYTGSELYLTALPDSPYKVCWV 480
QY 481 PMETSNLYLFDPTVCLETETAPLMYNTPTTLNREKEKPYKNPLPIAALIGGAVLV 540
Db 481 PMETSNLYLFDPTVCLETETAPLMYNTPTTLNREKEKPYKNPLPIAALIGGAVLV 540
QY 541 TIALALVYVYHNGSLFSRNCVSKGRRKDDYAEAGTKKNSILETRFSQMLPIS 600
Db 541 TIALALVYVYHNGSLFSRNCVSKGRRKDDYAEAGTKKNSILETRFSQMLPIS 600
QY 601 NEPISEKEFVTHITFPNGMNLKNNHSESSNSRYSRDSGIPDSHSHS 649
Db 601 NEPISEKEFVTHITFPNGMNLKNNHSESSNSRYSRDSGIPDSHSHS 649

RESULT 3
AAB70533
ID AAB70533 standard; Protein; 649 AA.
XX
AC AAB70533;
XX
DT 09-MAY-2001 (first entry)
XX
DE Human PRO3 protein sequence SEQ ID NO:6.
XX
KW Human; PRO; PROX; cytostatic; immunomodulatory; reproduction;
KW gene therapy; cell proliferation; differentiation disorder; cancer;
KW immune associated disorder; gestational disease; pre-clampsia;
KW chromosome 20.
XX
OS Homo sapiens.
XX
PN WO200110902-A2.
XX
PD 15-FEB-2001.
XX
PF 11-AUG-2000; 2000WO-US21857.
XX
PR 11-AUG-1999; 99US-0148433.
PR 10-AUG-2000; 2000US-0148433.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shinkets RA, Fernandes B;
XX
DR WPI; 2001-147509/15.
DR N-PSDB; AAF74434.
XX
PT Nucleic acids encoding secreted polypeptides, designated PROX
PT polypeptides, useful for treating a syndrome associated with a
PT PROX-associated disorder, e.g. cancer
XX
PS Claim 1; Page 13-15; 166pp; English.
XX
CC The present invention describes isolated nucleic acids encoding secreted
CC polypeptides, designated PROX polypeptides (i.e. a PRO polypeptide where
CC X is an integer from 1 to 17). PROX polypeptides have cytosolic,
CC immunomodulatory and reproduction activities, and can be used in gene
CC therapy, and as PROX antagonists and PROX agonists. PROX polypeptides,
CC nucleic acids and antibodies are useful in the manufacture of a
CC medicament for treating a syndrome associated with a PROX-associated
CC disorder, e.g. a cell proliferation and/or differentiation disorder
CC (e.g. cancer or immune associated disorders) and a gestational disease
CC (e.g. pre-clampsia). They are also used for screening for a modulator of
CC activity or of latency or predisposition to a PROX-associated disorder.
CC AAF74432 to AAF74438 encode the specifically claimed human PROX
CC polypeptides PRO1 to PRO17 given in AAB70531 to AAB70547.
XX
SQ Sequence 649 AA;
Query Match 100.0%; Score 3400; DB 22; Length 649;
Best Local Similarity 100.0%; Pred. No. 7.5e-290;
Matches 649; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MISAWSFLIGKIGLQVAPLSVMAKSCPSVCRDAGFIYCNDRFTSTPTGIPEDA 60
Db 1 MISAWSFLIGKIGLQVAPLSVMAKSCPSVCRDAGFIYCNDRFTSTPTGIPEDA 60
QY 61 TTYLQNNQINNAGIPSDKLNLLKVERIYLYHNSIDFPTNPKYKELHLOENIRIT 120
Db 61 TTYLQNNQINNAGIPSDKLNLLKVERIYLYHNSIDFPTNPKYKELHLOENIRIT 120
QY 121 YDSLSKIPYLEELHLDNDSVSAVSEIEGAFRDSNLYRLFLSRNHLSTIPWGLPTIEEL 180
Db 121 YDSLSKIPYLEELHLDNDSVSAVSEIEGAFRDSNLYRLFLSRNHLSTIPWGLPTIEEL 180
QY 181 RDDNRISTSSPSLOGLSLKRLVLDGNLLNHHGLDKFFNLVNLSTELSVRSLTAA 240

Applied

Db 181 RLDNRISTSSPLOGTSLKRLVLDGNLNNHGLGDKVFENLVNLFELSLVNSLTAA 240
QY 241 PVNLPCTNLRKLYLQDNHNRVPPNAPSYLRQLYRDLMSNNLSNLPQIGIEDLDNITQL 300
Db 241 PVNLPCTNLRKLYLQDNHNRVPPNAPSYLRQLYRDLMSNNLSNLPQIGIEDLDNITQL 300
QY 301 ILRNNPVCCKMKWYDQSLPQVKNVRLMCOAPEKVRGMAIKDLNAELFDCKDSGI 360
Db 301 ILRNNPVCCKMKWYDQSLPQVKNVRLMCOAPEKVRGMAIKDLNAELFDCKDSGI 360
QY 361 VSTQITAIPTNVPYPAQGNFAPYKQPDIKNPKTKQOQTGSPSKTITIVKSVTS 420
Db 361 VSTQITAIPTNVPYPAQGNFAPYKQPDIKNPKTKQOQTGSPSKTITIVKSVTS 420
QY 421 DTIHSKWLPMALRLSWLKGHSFAGSITETIVTGERSEYLVTALEPDPSPYKVMV 480
Db 421 DTIHSKWLPMALRLSWLKGHSFAGSITETIVTGERSEYLVTALEPDPSPYKVMV 480
QY 481 PMETSNIYLFDETPVCIEPTAPLRMNPNTTLNREQKEPYKNPLPLAALIGGAVLV 540
Db 481 PMETSNIYLFDETPVCIEPTAPLRMNPNTTLNREQKEPYKNPLPLAALIGGAVLV 540
QY 541 TIALALVCHVYHRNGSLFSRNCAYSKGRKKDDYAFAGTKKDNLSILIRTSFQMLPIS 600
Db 541 TIALALVCHVYHRNGSLFSRNCAYSKGRKKDDYAFAGTKKDNLSILIRTSFQMLPIS 600
QY 601 NEPIKSEEFVHIFPPNGMNLKNNHSESSNSRYSRDSGIPDSHSHS 649
Db 601 NEPIKSEEFVHIFPPNGMNLKNNHSESSNSRYSRDSGIPDSHSHS 649

RESULT 4
ABB95559
ID ABB95559 standard; Protein; 649 AA.
XX ABB95559;
AC ABB95559;
DT 19-JUL-2002 (first entry)
XX Human angiogenesis related protein PRO1865 SEQ ID NO: 274.
XX Human; angiogenesis; PRO protein; cardiovascularisation; wound; cancer;
KW atherosclerosis; cardiac hypertrophy; gene therapy; endothelial disorder;
KW cardiant; cytostatic; antiangiogenic; hypotensive; vulnary;
XX antiarteriosclerotic.
XX Homo sapiens.
XX WO200208284-A2.
XX 31-JAN-2002.
XX 09-JUL-2001; 2001WO-US21735.
XX 20-JUL-2000; 2000US-219556P.
XX 25-JUL-2000; 2000US-220624P.
XX 25-JUL-2000; 2000US-220664P.
XX 28-JUL-2000; 2000WO-US20710.
XX 02-AUG-2000; 2000US-222695P.
XX 17-AUG-2000; 2000US-0643657.
XX 23-AUG-2000; 2000WO-US23522.
XX 24-AUG-2000; 2000WO-US23328.
XX 07-SEP-2000; 2000US-230978P.
XX 15-SEP-2000; 2000US-000000P.
XX 18-SEP-2000; 2000US-0664610.
XX 18-SEP-2000; 2000US-0665350.
XX 24-OCT-2000; 2000US-242922P.
XX 08-NOV-2000; 2000US-0709238.
XX 08-NOV-2000; 2000WO-US30952.
XX 10-NOV-2000; 2000WO-US30873.
XX 01-DEC-2000; 2000WO-US32678.
XX 20-DEC-2000; 2000US-0747259.

Handwritten signature

20-DEC-2000; 2000WO-US34956.
22-JAN-2001; 2001US-0767609.
28-FEB-2001; 2001US-0796498.
28-FEB-2001; 2001WO-US06520.
01-MAR-2001; 2001WO-US06666.
09-MAR-2001; 2001US-0802706.
14-MAR-2001; 2001US-0808689.
22-MAR-2001; 2001US-0816744.
05-APR-2001; 2001US-0828366.
10-MAY-2001; 2001US-0854208.
10-MAY-2001; 2001US-0854280.
25-MAY-2001; 2001US-0866028.
25-MAY-2001; 2001US-0866034.
25-MAY-2001; 2001WO-US17092.
30-MAY-2001; 2001US-0870574.
30-MAY-2001; 2001WO-US17443.
01-JUN-2001; 2001WO-US17800.
20-JUN-2001; 2001WO-US19692.
28-JUN-2001; 2001WO-US00000.
(GETH) GENENTECH INC.
(BAKE) BAKER K P.
(FERR) FERRARA N.
(GERB) GERBER H.
(GERR) GERRITSEN M E.
(GODD) GODDARD A.
(GODO) GODOWSKI P J.
(GURN) GURNEY A L.
(HILL) HILLAN K J.
(MARS) MARSTERS S A.
(PANJ) PAN J.
(PAON) PAONI N F.
(STEP) STEPHAN J F.
(WATA) WATANABE C K.
(WILL) WILLIAMS P M.
(WOOD) WOOD W I.
Baker KP, Ferrara N, Gerber H, Gerritsen ME, Goddard A;
Godowski PJ, Gurney AL, Hillan KJ, Marsters SA, Pan J, Paoni NF;
Stephan JF, Watanabe CK, Williams PM, Wood WI, Ye W;
WPI: 2002-171999/22.
N-PSDB: ABL95697.
One hundred and eighty seven nucleic acids encoding PRO polypeptides,
useful in diagnosis and treatment of cardiovascular (e.g. myocardial
infarction), endothelial or angiogenic disorders in a mammal -
Claim 11; Fig 274; 567pp; English.
The present invention provides the protein and coding sequences of human
PRO proteins. These are useful for treating or diagnosing a
cardiovascular, endothelial or angiogenic disorder, including cardiac
hypertrophy, trauma, cancer, age-related macular degeneration,
atherosclerosis, hypertension, arterial restenosis, rheumatoid arthritis,
angina, myocardial infarctions, thrombophlebitis, lymphangitis, tumour
angiogenesis (such as breast carcinoma and liver carcinoma) and wound
healing. The present sequence is a PRO protein of the invention.
XX Sequence 649 AA;
Query Match 100.0%; Score 3400; DB 23; Length 649;
Best Local Similarity 100.0%; Pred. No. 7.5e-290;
Matches 649; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MTSAAWSIFLIGTKIGLEQVAPLSVMAKSCPSVCRDAGFYCNDRLTSTPTGIPEDA 60
Db 1 MTSAAWSIFLIGTKIGLEQVAPLSVMAKSCPSVCRDAGFYCNDRLTSTPTGIPEDA 60
QY 61 TTYLQNNQINNAGIPSDKLNLLKVERIYLYHNSIDEFPNLPKYVKELHQLQNNITIT 120
Db 61 TTYLQNNQINNAGIPSDKLNLLKVERIYLYHNSIDEFPNLPKYVKELHQLQNNITIT 120

```
QY 121 YDLSKIPYLEELHLDNSVSASVIEGAFRDSNLYRLLEFISRNHLSTIPWGLPTIEEL 180
Db 121 YDLSKIPYLEELHLDNSVSASVIEGAFRDSNLYRLLEFISRNHLSTIPWGLPTIEEL 180
QY 181 RLDNRISTISSPSLGTSIKELVDGNLNNHGLGDKVFFNLVNLTELSVRSNLSFAA 240
Db 181 RLDNRISTISSPSLGTSIKELVDGNLNNHGLGDKVFFNLVNLTELSVRSNLSFAA 240
QY 241 PYNLPGTNLRKLYLQDNEHNRVPPNFAFSLRYLRDMSNNLSNLPQGFDDLDNIQTOL 300
Db 241 PYNLPGTNLRKLYLQDNEHNRVPPNFAFSLRYLRDMSNNLSNLPQGFDDLDNIQTOL 300
QY 301 ILRNNPWCCKMKWMDWLOSLPVKVVNVRGLMCAPEKVRGMAIKDLNAELFDCCKDSGI 360
Db 301 ILRNNPWCCKMKWMDWLOSLPVKVVNVRGLMCAPEKVRGMAIKDLNAELFDCCKDSGI 360
QY 361 VSTIQTITAIPTNVTYPAQGWPAVTKQPDINKPKLRDQQTGSPSRKTIITVKSVTS 420
Db 361 VSTIQTITAIPTNVTYPAQGWPAVTKQPDINKPKLRDQQTGSPSRKTIITVKSVTS 420
QY 421 DTIHISWKLALPMTALRLSLKLGHPAFGSTTETIVTGERSEYLVTALEPDSPIKVCMV 480
Db 421 DTIHISWKLALPMTALRLSLKLGHPAFGSTTETIVTGERSEYLVTALEPDSPIKVCMV 480
QY 481 PMETSNLYLDETVCYTETAPLRMYNPTTLNRQKEPKYKPNPLPLAALIGGAVALV 540
Db 481 PMETSNLYLDETVCYTETAPLRMYNPTTLNRQKEPKYKPNPLPLAALIGGAVALV 540
QY 541 TIALALVGVVYHRRGSLFRNCAYSKGRRRDKDYAEAGTKKNSILEIRETSFQMLPIS 600
Db 541 TIALALVGVVYHRRGSLFRNCAYSKGRRRDKDYAEAGTKKNSILEIRETSFQMLPIS 600
QY 601 NEPISEKEEVIHTIFFPPNGMNLKNNHSESSNRSYRDSGIPDSHSHS 649
Db 601 NEPISEKEEVIHTIFFPPNGMNLKNNHSESSNRSYRDSGIPDSHSHS 649

RESULT 5
ABB84953
ID ABB84953 standard; Protein: 649 AA.
XX AC ABB84953;
XX DT 16-MAY-2002 (first entry)
XX DE Human PRO1865 protein sequence SEQ ID NO:274.
XX KW Human; angiogenesis; cardiant; cytostatic; antiangiogenic; hypotensive;
XX KW vulnery; antiarteriosclerotic; PRO agonist; PRO antagonist; trauma;
XX KW gene therapy; cardiovascular disorder; endothelial disorder; cancer;
XX KW angiogenic disorder; cardiac hypertrophy; atherosclerosis; hypertension;
XX KW age-related macular degeneration; arterial restenosis; angina;
XX KW rheumatoid arthritis; myocardial infarction; thrombophlebitis;
XX KW lymphangitis; tumour angiogenesis; breast carcinoma; liver carcinoma;
XX KW wound healing; chromosome mapping; gene mapping.
XX OS Homo sapiens.
XX FN WC200200690-A2.
XX PD 03-JAN-2002.
XX PF 20-JUN-2001; 2001WO-US19692.
XX PR 23-JUN-2000; 2000US-213637P.
XX PR 20-JUL-2000; 2000US-219556P.
XX PR 25-JUL-2000; 2000US-220624P.
XX PR 28-JUL-2000; 2000US-220664P.
XX PR 02-AUG-2000; 2000WO-US20710.
XX PR 17-AUG-2000; 2000US-222695P.
XX PR 23-AUG-2000; 2000US-0643657.
XX PR 24-AUG-2000; 2000WO-US23522.
XX PR 24-AUG-2000; 2000WO-US23328.
```

```
PR 07-SEP-2000; 2000US-230978P.
PR 18-SEP-2000; 2000US-0664610.
PR 18-SEP-2000; 2000US-0665350.
PR 24-OCT-2000; 2000US-242922P.
PR 08-NOV-2000; 2000US-0709238.
PR 08-NOV-2000; 2000WO-US30952.
PR 10-NOV-2000; 2000WO-US30873.
PR 01-DEC-2000; 2000WO-US32678.
PR 20-DEC-2000; 2000US-0747259.
PR 20-DEC-2000; 2000WO-US34956.
PR 22-JAN-2001; 2001US-0767609.
PR 28-FEB-2001; 2001US-0796498.
PR 28-FEB-2001; 2001WO-US06520.
PR 01-MAR-2001; 2001WO-US06666.
PR 09-MAR-2001; 2001US-0802706.
PR 14-MAR-2001; 2001US-0808689.
PR 22-MAR-2001; 2001US-0816744.
PR 05-APR-2001; 2001US-0828366.
PR 10-MAY-2001; 2001US-0854208.
PR 10-MAY-2001; 2001US-0854280.
PR 25-MAY-2001; 2001US-0866028.
PR 25-MAY-2001; 2001US-0866034.
PR 25-MAY-2001; 2001WO-US17092.
PR 30-MAY-2001; 2001US-0870574.
PR 30-MAY-2001; 2001WO-US17443.
PR 01-JUN-2001; 2001WO-US17800.
XX (GETH ) GENENTECH INC.
XX PA Baker KP, Ferrara N, Gerber H, Gerritsen ME, Goddard A;
XX PI Godowski PJ, Gurney AL, Hillan KJ, Marsters SA, Pan J, Paoni NF;
XX PI Stephan JF, Watanabe CK, Williams PM, Wood WI, Ye W;
XX DR WPI; 2002-090516/12.
XX DR N-PSDB; ABL88208.
XX XX
XX PT One hundred and eighty seven nucleic acids encoding PRO polypeptides,
XX PT useful in diagnosis and treatment of cardiovascular (e.g. myocardial
XX PT infarction), endothelial or angiogenic disorders in a mammal -
XX PS Claim 11; Fig 274; 565pp; English.
XX CC ABL88072 to ABL88258 encode the PRO proteins given in ABB84817 to
XX CC ABB85003. The PRO proteins and polynucleotides have cardiant, cytostatic,
XX CC antiangiogenic, hypotensive, vulnery and antiarteriosclerotic
XX CC activities, and can be used in gene therapy. The PRO polynucleotides,
XX CC proteins, agonists and antagonists are useful for treating or diagnosing
XX CC a cardiovascular, endothelial or angiogenic disorder in a mammal,
XX CC e.g. cardiac hypertrophy, trauma, cancer, age-related macular
XX CC degeneration, atherosclerosis, hypertension, arterial restenosis,
XX CC rheumatoid arthritis, angina, myocardial infarctions, thrombophlebitis,
XX CC lymphangitis, tumour angiogenesis (such as breast carcinoma and liver
XX CC carcinoma) and wound healing. The PRO polynucleotides have applications
XX CC in molecular biology, including use as hybridisation probes, and in
XX CC chromosome and gene mapping. ABL88259 to ABL88267 represent primers and
XX CC probes used in the exemplification of the present invention.
XX SQ Sequence 649 AA;
XX Query Match 100.0%; Score 3400; DB 23; Length 649;
XX Best Local Similarity 100.0%; Pred. No. 7.5e-290;
XX Matches 649; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MISAAWSIFLIGTKIGLFLQVAPLSVMAKSCPSVCRDAGFIYCNDRFLSIFPTGIPEDA 60
Db 1 MISAAWSIFLIGTKIGLFLQVAPLSVMAKSCPSVCRDAGFIYCNDRFLSIFPTGIPEDA 60
QY 61 TTYLQNNQINNAGIPSDLKNLKAVERIYLYHNSLDEPTNLPKYVELHLQENNRITIT 120
Db 61 TTYLQNNQINNAGIPSDLKNLKAVERIYLYHNSLDEPTNLPKYVELHLQENNRITIT 120
QY 121 YDSISKIPYLEELHLDNSVSASVIEGAFRDSNLYRLLEFISRNHLSTIPWGLPTIEEL 180
XX 121 YDSISKIPYLEELHLDNSVSASVIEGAFRDSNLYRLLEFISRNHLSTIPWGLPTIEEL 180
```

Db 121 YDSLSKIPYLEELHLDNDSVAVSIEGAFRDSNYLRLLFLSRNHLSTIPWGLPRTTEEL 180

QY 181 RLDDNRISTSSPSLQGLTSLKRLVLDGNLNNHGLGDKVFNLYNTELSLVNSLTA 240

Db 181 RLDDNRISTSSPSLQGLTSLKRLVLDGNLNNHGLGDKVFNLYNTELSLVNSLTA 240

QY 241 PVNLPGNLRLKLYLDQNHINRVPNAPSYLRQLYRLDMSNNLSNLPQIFDDLDNTOL 300

Db 241 PVNLPGNLRLKLYLDQNHINRVPNAPSYLRQLYRLDMSNNLSNLPQIFDDLDNTOL 300

QY 301 ILRNNPWCCKMKWVRLQSLPLVKNVRLGMCQAPKVRGMAIKDLNLELFDCKDSGI 360

Db 301 ILRNNPWCCKMKWVRLQSLPLVKNVRLGMCQAPKVRGMAIKDLNLELFDCKDSGI 360

QY 361 VSTIQTITAPNTVPAQOWPAPVTKQPKDKNPKLTQDQQTGSPSKTITIVKSVTS 420

Db 361 VSTIQTITAPNTVPAQOWPAPVTKQPKDKNPKLTQDQQTGSPSKTITIVKSVTS 420

QY 421 DTIHSWKALPMTALRLSWLKLGHSPAFGSITETIVTGERSEYLVTALEPDSPIKVCW 480

Db 421 DTIHSWKALPMTALRLSWLKLGHSPAFGSITETIVTGERSEYLVTALEPDSPIKVCW 480

QY 481 PMETSNIYLFDETPVCIEETAPLMYNTPTTLNRQEKPKYKNNPLAALIGGAVLV 540

Db 481 PMETSNIYLFDETPVCIEETAPLMYNTPTTLNRQEKPKYKNNPLAALIGGAVLV 540

QY 541 TIALALVWVHRNGSLFSRNCAYSKGRKKDDYAGTCKDKNLSILIRETSFQMLPIS 600

Db 541 TIALALVWVHRNGSLFSRNCAYSKGRKKDDYAGTCKDKNLSILIRETSFQMLPIS 600

QY 601 NEPISKEEFVHIHIFPPNGMNLKNNHSESSNSRYSRDSGIPDSDHS 649

Db 601 NEPISKEEFVHIHIFPPNGMNLKNNHSESSNSRYSRDSGIPDSDHS 649

RESULT 6

AAB93758

ID AAB93758 standard; Protein; 649 AA.

XX AAB93758;

XX 26-JUN-2001 (first entry)

XX Human protein sequence SEQ ID NO:13435.

XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.

XX Homo sapiens.

XX EP1074617-A2.

XX 07-FEB-2001.

XX 28-JUL-2000; 2000EP-0116126.

XX 29-JUL-1999; 99JP-0248036.

XX 27-AUG-1999; 99JP-0300253.

XX 11-JAN-2000; 2000JP-0118776.

XX 02-MAY-2000; 2000JP-0183767.

XX 09-JUN-2000; 2000JP-0241899.

XX (HELI-) HELIX RES INST.

XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX WPI; 2001-318749/34.

XX Primer sets for synthesizing polynucleotides, particularly the 5602

PT full-length cDNAs defined in the specification, and for the detection

PT and/or diagnosis of the abnormality of the proteins encoded by the

PT full-length cDNAs -

XX

PS Claim 8; SEQ ID 13435; 2537bp + CD ROM; English.

XX The present invention describes primer sets for synthesizing 5602

CC full-length cDNAs defined in the specification. Where a primer set

CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary

CC to the complementary strand of a polynucleotide which comprises one of

CC the 5602 nucleotide sequences defined in the specification, where the

CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination

CC of an oligonucleotide comprising a sequence complementary to the

CC complementary strand of a polynucleotide which comprises a 5'-end

CC sequence and an oligonucleotide comprising a sequence complementary to a

CC polynucleotide which comprises a 3'-end sequence, where the

CC oligonucleotide comprises at least 15 nucleotides and the combination of

CC the 5'-end sequence/3'-end sequence is selected from those defined in

CC the specification. The primer sets can be used in antisense therapy and

CC in gene therapy. The primers are useful for synthesizing polynucleotides,

CC particularly full-length cDNAs. The primers are also useful for the

CC detection and/or diagnosis of the abnormality of the proteins encoded by

CC the full-length cDNAs. The primers allow obtaining of the full-length

CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and

CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to

CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632

CC represent oligonucleotides, all of which are used in the exemplification

CC of the present invention.

XX Sequence 649 AA:

SQ Query Match 99.9%; Score 3396; DB 22; Length 649;

Best Local Similarity 99.8%; Pred. No. 1.7e-289;

Matches 648; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MISAAWSIFLIGTKIGLFLQVAPLSVMAKSCPSVCRDAGFYCNDRLTISPTIGPEDA 60

Db 1 MISAAWSIFLIGTKIGLFLQVAPLSVMAKSCPSVCRDAGFYCNDRLTISPTIGPEDA 60

QY 61 TTYLQNNQNNAGIPSDLKLLKVERIYLYHNSLDEFFTNLPKYVHELHQLQENNTIT 120

Db 61 TTYLQNNQNNAGIPSDLKLLKVERIYLYHNSLDEFFTNLPKYVHELHQLQENNTIT 120

QY 121 YDSLSKIPYLEELHLDNDSVAVSIEGAFRDSNYLRLLFLSRNHLSTIPWGLPRTTEEL 180

Db 121 YDSLSKIPYLEELHLDNDSVAVSIEGAFRDSNYLRLLFLSRNHLSTIPWGLPRTTEEL 180

QY 181 RLDDNRISTSSPSLQGLTSLKRLVLDGNLNNHGLGDKVFNLYNTELSLVNSLTA 240

Db 181 RLDDNRISTSSPSLQGLTSLKRLVLDGNLNNHGLGDKVFNLYNTELSLVNSLTA 240

QY 241 PVNLPGNLRLKLYLDQNHINRVPNAPSYLRQLYRLDMSNNLSNLPQIFDDLDNTOL 300

Db 241 PVNLPGNLRLKLYLDQNHINRVPNAPSYLRQLYRLDMSNNLSNLPQIFDDLDNTOL 300

QY 301 ILRNNPWCCKMKWVRLQSLPLVKNVRLGMCQAPKVRGMAIKDLNLELFDCKDSGI 360

Db 301 ILRNNPWCCKMKWVRLQSLPLVKNVRLGMCQAPKVRGMAIKDLNLELFDCKDSGI 360

QY 361 VSTIQTITAPNTVPAQOWPAPVTKQPKDKNPKLTQDQQTGSPSKTITIVKSVTS 420

Db 361 VSTIQTITAPNTVPAQOWPAPVTKQPKDKNPKLTQDQQTGSPSKTITIVKSVTS 420

QY 421 DTIHSWKALPMTALRLSWLKLGHSPAFGSITETIVTGERSEYLVTALEPDSPIKVCW 480

Db 421 DTIHSWKALPMTALRLSWLKLGHSPAFGSITETIVTGERSEYLVTALEPDSPIKVCW 480

QY 481 PMETSNIYLFDETPVCIEETAPLMYNTPTTLNRQEKPKYKNNPLAALIGGAVLV 540

Db 481 PMETSNIYLFDETPVCIEETAPLMYNTPTTLNRQEKPKYKNNPLAALIGGAVLV 540

QY 541 TIALALVWVHRNGSLFSRNCAYSKGRKKDDYAGTCKDKNLSILIRETSFQMLPIS 600

Db 541 TIALALVWVHRNGSLFSRNCAYSKGRKKDDYAGTCKDKNLSILIRETSFQMLPIS 600

QY 601 NEPISKEEFVHIHIFPPNGMNLKNNHSESSNSRYSRDSGIPDSDHS 649

Db 601 NEPISKEEFVHIHIFPPNGMNLKNNHSESSNSRYSRDSGIPDSDHS 649

Db 601 NEPISKEEFVHTIFFPENGMLYKNNHSESSNSRYSRDSGIPDSHSHS 649

RESULT 7

AAW78907

XX AC AAW78907; standard; Protein; 649 AA.

XX DT 06-NOV-2001 (first entry)

XX DE Human protein SEQ ID NO 1569.

XX KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;

XX KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;

XX KW tissue growth factor; immunomodulatory; cancer; leukaemia;

XX KW nervous system disorder; arthritis; inflammation.

XX OS Homo sapiens.

XX PN WO200157190-A2.

XX PD 09-AUG-2001.

XX PF 05-FEB-2001; 2001WO-US04098.

XX PR 03-FEB-2000; 2000US-0496914.

XX PR 27-APR-2000; 2000US-0560875.

XX PR 20-JUN-2000; 2000US-0598075.

XX PR 19-JUL-2000; 2000US-0620325.

XX PR 01-SEP-2000; 2000US-0654936.

XX PR 15-SEP-2000; 2000US-0663561.

XX PR 20-OCT-2000; 2000US-0693325.

XX PR 30-NOV-2000; 2000US-0728422.

XX PA (HYSE-) HYSEQ INC.

XX PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;

XX PI Zhao QH, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;

XX PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;

XX WPI: 2001-476283/51.

XX N-PSDB; AAK52040.

XX PT Nucleic acids encoding polypeptides with cytokine-like activities,

XX PT useful in diagnosis and gene therapy -

XX PS Claim 20; Page 3895-3896; 6221pp; English.

XX CC The invention relates to polynucleotides (AAK51456-AAK53435) and the

XX CC encoded polypeptides (AAW78323-AAW80302) that exhibit activity relating to

XX CC cytokine, cell proliferation or cell differentiation or which may induce

XX CC production of other cytokines in other cell populations. The

XX CC polynucleotides and polypeptides are useful in gene therapy, vaccines or

XX CC peptide therapy. The polypeptides have various cytokine-like activities,

XX CC e.g. stem cell growth factor activity, haematopoiesis regulating

XX CC activity, tissue growth factor activity, immunomodulatory activity and

XX CC activator/inhibitor activity and may be useful in the diagnosis and/or

XX CC treatment of cancer, leukaemia, nervous system disorders, arthritis and

XX CC inflammation.

XX CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666

XX CC (AAW80020) are omitted as the relevant pages from the sequence listing

XX CC were missing at the time of publication.

XX SQ Sequence 649 AA;

Query Match 99.9%; Score 3395; DB 22; Length 649;

Best Local Similarity 99.8%; Pred. No. 2.le-289;

Matches 648; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MISAWSIFLGTIGLFLQVAPLSVMAKSCPSVCRDAGFYCNDRFYTSIPTGIPEDA 60

Db 1 MISAWSIFLGTIGLFLQVAPLSVMAKSCPSVCRDAGFYCNDRFYTSIPTGIPEDA 60

QY 61 TTYLQNNQINNAGIPSDLNLLKVERIYVHNSIDDEFTNLPKYVKELHLOENNIRIT 120

Db 61 TTYLQNNQINNAGIPSDLNLLKVERIYVHNSIDDEFTNLPKYVKELHLOENNIRIT 120

QY 121 YDSLKIPYLEELHLDNDSVSAVIEGAFRDSNYLRLLFLSRNHLSTIPWGLPTIEEL 180

Db 121 YDSLKIPYLEELHLDNDSVSAVIEGAFRDSNYLRLLFLSRNHLSTIPWGLPTIEEL 180

QY 181 RLDDNRISTISSPSLOGLTSLKRLVLDGNLNNHGLGDKYFFNLVNLTELSVRSNTAA 240

Db 181 RLDDNRISTISSPSLOGLTSLKRLVLDGNLNNHGLGDKYFFNLVNLTELSVRSNTAA 240

QY 241 PVNLPSTNLRKLYLDNHNIRVPNPAFSTLRQLYRLDMSNNLSNLPQGFIDLDNITOL 300

Db 241 PVNLPSTNLRKLYLDNHNIRVPNPAFSTLRQLYRLDMSNNLSNLPQGFIDLDNITOL 300

QY 301 ILRNPWTGCKMKWVRDMLQSLPVKVNVRGLMQCAPEKVRGMAIKDLNAELFCKDSGI 360

Db 301 ILRNPWTGCKMKWVRDMLQSLPVKVNVRGLMQCAPEKVRGMAIKDLNAELFCKDSGI 360

QY 361 VSTIQTITTAIPNTVYPAQGWPAVTKQPDINKPKLTKDOOTGSPSRKTITITVKSPTS 420

Db 361 VSTIQTITTAIPNTVYPAQGWPAVTKQPDINKPKLTKDOOTGSPSRKTITITVKSPTS 420

QY 421 DTIHISWKLALPMTALRLSWLKLHSPAFGSTTETITVTSRSEYLVTALEPDSYKVCMV 480

Db 421 DTIHISWKLALPMTALRLSWLKLHSPAFGSTTETITVTSRSEYLVTALEPDSYKVCMV 480

QY 481 PMETSNIYLFDETPVCIETETAPLRMYPTTLARQEKPEYKPNPLPLAALIGGVALV 540

Db 481 PMETSNIYLFDETPVCIETETAPLRMYPTTLARQEKPEYKPNPLPLAALIGGVALV 540

QY 541 TIALALVGVTVHNGSLFRNCAYSKGRKRDYAEAGTKKDNSTLETRFSOMLPI 600

Db 541 TIALALVGVTVHNGSLFRNCAYSKGRKRDYAEAGTKKDNSTLETRFSOMLPI 600

QY 601 NEPISKEEFVHTIFFPENGMLYKNNHSESSNSRYSRDSGIPDSHSHS 649

Db 601 NEPISKEEFVHTIFFPENGMLYKNNHSESSNSRYSRDSGIPDSHSHS 649

RESULT 8

AAW93408

XX ID AAW93408 standard; Protein; 649 AA.

XX AC AAW93408;

XX DT 06-NOV-2001 (first entry)

XX DE Human polypeptide, SEQ ID NO: 3016.

XX KW Human; full length cDNA; cDNA synthesis; oligo-capping.

XX OS Homo sapiens.

XX PN EP1130094-A2.

XX PD 05-SEP-2001.

XX PF 07-JUL-2000; 2000EP-0114089.

XX PR 08-JUL-1999; 99JP-0194486.

XX PR 11-JAN-2000; 2000JP-0118774.

XX PR 02-MAY-2000; 2000JP-0183765.

XX PA (HELI-) HELIX RES INST.

XX PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;

XX PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;

XX DR WPI: 2001-524255/58.

XX N-PSDB; AAK94329.

XX 830 Primers useful for synthesizing full length cDNA clones and their
 PT use in genetic manipulation -
 XX
 XX Claim 8; SEQ ID NO 3016; 1380pp + sequence listing; English.
 XX
 CC The invention relates to primers for synthesizing full length cDNA
 CC clones. 830 cDNA molecules encoding a human protein have been
 CC isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA
 CC molecules have been determined. Primers for synthesizing the full length
 CC cDNA are useful for clarifying the function of the protein encoded by
 CC the cDNA. The full length clones were obtained by construction of full,
 CC length enriched cDNA libraries that were synthesised by the oligo-capping
 CC method. The primers enable the production of the full length cDNA easily
 CC without any special methods. The present sequence is a polypeptide
 CC encoded by a full length human cDNA of the invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in CD-ROM format directly from EPO.
 XX
 XX Sequence 649 AA;

Query Match 99.9%; Score 3395; DB 22; Length 649;
 Best Local Similarity 99.8%; Pred. No. 2.1e-289;
 Matches 648; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MISAWSIFLIGTKIGLFTQVAPLSVMAKSCPSVCRDAGFYCNDRTFTSIPTGIPEDA 60
 Db 1 MISAWSIFLIGTKIGLFTQVAPLSVMAKSCPSVCRDAGFYCNDRTFTSIPTGIPEDA 60
 QY 61 TTYLQNNQINNAIGPSDLKLLKVERIYLYHNSLDEFFTNLPKYKELHLOENNITIT 120
 Db 61 TTYLQNNQINNAIGPSDLKLLKVERIYLYHNSLDEFFTNLPKYKELHLOENNITIT 120
 QY 121 YDSLKIPYLEELHLDNDSVSAVSEEGAFRDSNYLRLFLSLNHLSTIPWGLPRTIEEL 180
 Db 121 YDSLKIPYLEELHLDNDSVSAVSEEGAFRDSNYLRLFLSLNHLSTIPWGLPRTIEEL 180
 QY 181 RLDNDRISTSSPSLGLSLKRLVLDGNLNNHGLGDKVFFNLVNLSTLSVRSNTAA 240
 Db 181 RLDNDRISTSSPSLGLSLKRLVLDGNLNNHGLGDKVFFNLVNLSTLSVRSNTAA 240
 QY 241 PVNLPFTNLRKLYLDNHNRYPPNAPSYLRQLYRLDMSNNLSNLPQIGIFDDLDNTQL 300
 Db 241 PVNLPFTNLRKLYLDNHNRYPPNAPSYLRQLYRLDMSNNLSNLPQIGIFDDLDNTQL 300
 QY 301 ILRNNPWCCKMKWRDMLQSLPVKVNVRGIMQCAPEKVRGMAIKDLNAELFCKDSGI 360
 Db 301 ILRNNPWCCKMKWRDMLQSLPVKVNVRGIMQCAPEKVRGMAIKDLNAELFCKDSGI 360
 QY 361 VSTIQTITAPNTVYPAQGWAPVTKPDINKPKLKDOQTGSPSRKTIITIVKSVTS 420
 Db 361 VSTIQTITAPNTVYPAQGWAPVTKPDINKPKLKDOQTGSPSRKTIITIVKSVTS 420
 QY 421 DTIHLNWKALPMTALRLSWLKLHSPAFGSITITIVTGERSELYLTALPDSPKYKCMV 480
 Db 421 DTIHLNWKALPMTALRLSWLKLHSPAFGSITITIVTGERSELYLTALPDSPKYKCMV 480
 QY 481 PMETSNLYLDEFPVCIETETAPLRMYNPTTLNREQEKEPYKNPPLAAIIGGAVLV 540
 Db 481 PMETSNLYLDEFPVCIETETAPLRMYNPTTLNREQEKEPYKNPPLAAIIGGAVLV 540
 QY 541 TIALALVCWYVHRNGLSRNCAYSKGRRKDDYAEAGTKDONSILETRTSFQMLPIS 600
 Db 541 TIALALVCWYVHRNGLSRNCAYSKGRRKDDYAEAGTKDONSILETRTSFQMLPIS 600
 QY 601 NEPISEKEEFVHTTFFPPNGNLYLKNHSESSNSRSYRDSGIPDSHSHS 649
 Db 601 NEPISEKEEFVHTTFFPPNGNLYLKNHSESSNSRSYRDSGIPDSHSHS 649
 RESULT 9
 AAY76050
 ID AAY76050 standard; Protein; 649 AA.

XX AAY76050;
 XX 27-MAR-2000 (first entry)
 XX Murine skin cell protein, SEQ ID NO:305.
 XX
 KW Skin; dermal papilla; keratinocyte; neonatal foreskin fibroblast;
 KW embryonic skin cell; keratinocyte stem cell; transit amplifying cell;
 KW secreted; transmembrane; inflammation; cancer; neurological disease;
 KW angiogenesis; tumour vascularisation; growth disorder;
 KW developmental disorder; skin wound; hair follicle disorder;
 KW anti-inflammatory; cytostatic; neuroprotective; vulnery.
 OS Mus sp.
 XX WO9955865-A1.
 XX 04-NOV-1999 -
 XX 29-APR-1999; 99WO-NZ00051.
 XX 29-APR-1998; 98US-0069726.
 XX 09-NOV-1998; 98US-0188930.
 XX (GENE-) GENESIS RES & DEV CORP LTD.
 XX Strachan L, Sleeman M, Watson JD, Onrust R, Kumble A, Murison JG;
 DR WPI; 2000-072117/06.
 DR N-PSDB; AAZ61755.
 XX Novel polynucleotides useful for the treatment of various conditions
 PT including wounds and cancer -
 XX Claim 4; Page 180-182; 235pp; English.
 CC The invention relates to novel nucleic acid sequences derived from rat
 CC dermal papilla, human keratinocytes and neonatal foreskin fibroblasts,
 CC and mouse embryonic skin, keratinocyte stem cells and transit amplifying
 CC cells. Polypeptides of the invention may be used to treat inflammation,
 CC cancer and neurological diseases. The proteins may be used to stimulate
 CC the growth and motility of keratinocytes, to inhibit the growth of
 CC cancer cells, to modulate angiogenesis and tumour vascularisation, to
 CC modulate skin inflammation, to modulate epithelial cell growth and to
 CC inhibit binding of HIV-1 to leukocytes. The invention may also be used
 CC to treat growth and developmental defects, skin wounds and hair follicle
 CC disorders. Sequences AAY75942-Y76123 represent polypeptides encoded
 CC by cDNA sequences derived from several mouse, rat or human skin cell
 CC types. Sequences AAY75942-Y75947, AAY76020-Y76021, AAY76094-Y76104 and
 CC AAY76119 are proteins with an N-terminal signal sequence, indicating
 CC that they are secreted. Sequences AAY75986-Y75989, AAY76061-Y76071,
 CC AAY76106-Y76109 and AAY76121-Y76122 are proteins with one or more
 CC putative transmembrane domains.
 XX Sequence 649 AA;
 SQ
 Query Match 96.6%; Score 3284; DB 21; Length 649;
 Best Local Similarity 96.5%; Pred. No. 1.2e-279;
 Matches 626; Conservative 11; Mismatches 12; Indels 0; Gaps 0;
 QY 1 MISAWSIFLIGTKIGLFTQVAPLSVMAKSCPSVCRDAGFYCNDRTFTSIPTGIPEDA 60
 Db 1 MISAWSIFLIGTKIGLFTQVAPLSVMAKSCPSVCRDAGFYCNDRTFTSIPTGIPEDA 60
 QY 61 TTYLQNNQINNAIGPSDLKLLKVERIYLYHNSLDEFFTNLPKYKELHLOENNITIT 120
 Db 61 TTYLQNNQINNAIGPSDLKLLKVERIYLYHNSLDEFFTNLPKYKELHLOENNITIT 120
 QY 121 YDSLKIPYLEELHLDNDSVSAVSEEGAFRDSNYLRLFLSLNHLSTIPWGLPRTIEEL 180
 Db 121 YDSLKIPYLEELHLDNDSVSAVSEEGAFRDSNYLRLFLSLNHLSTIPWGLPRTIEEL 180

bad date

QY 181 RLDNRISTISSPSLOGLTSLKRLVLDGNLNNHGLGDKVFNFLVNLTELSELVRSNLTA 240
 Db RLDNRISTISSPSLOGLTSLKRLVLDGNLNNHGLGDKVFNFLVNLTELSELVRSNLTA 240
 QY 241 PVNLPOTNLRLKLYLDQNHINRVPNFAFSLRYLYRLDMSNNLSNLPQGFDDLDNITQL 300
 Db PVNLPOTNLRLKLYLDQNHINRVPNFAFSLRYLYRLDMSNNLSNLPQGFDDLDNITQL 300
 QY 301 ILRNNPWYCGCKMKWVRDLQSLPVKNVVRGLMCOAPEKVRGMAIKDLNAELFDCKDSGI 360
 Db ILRNNPWYCGCKMKWVRDLQSLPVKNVVRGLMCOAPEKVRGMAIKDLNAELFDCKDSGI 360
 QY 361 VSTIQTITPAINTVYPAQGWPAVTKQPDIKNPKLTQDQTTGSPSRKTIITVKSPTS 420
 Db VSTIQTITPAINTVYPAQGWPAVTKQPDIKNPKLTQDQTTGSPSRKTIITVKSPTS 420
 QY 421 DTIHISWALPMTALRLSLWKLGHSPAFGSIETIVTGERSEYIVTALPESPVKCV 480
 Db DTIHISWALPMTALRLSLWKLGHSPAFGSIETIVTGERSEYIVTALPESPVKCV 480
 QY 481 PMETSNLYLDFETPVCIEETAPLRMTNPTTLNRQEKPEYKKNPNPLAAIIGGVALV 540
 Db PMETSNLYLDFETPVCIEETAPLRMTNPTTLNRQEKPEYKKNPNPLAAIIGGVALV 540
 QY 541 TIALALVGVYHRNGSLFSRNCAYSKGRRKDDYAEAGTKKDNSILEIRETSFQMLPIS 600
 Db TIALALVGVYHRNGSLFSRNCAYSKGRRKDDYAEAGTKKDNSILEIRETSFQMLPIS 600
 QY 601 NEPISKEEFVIHTIPPPNGMNLKNNHSESSNRSYRDSGIPDSDSHS 649
 Db NEPISKEEFVIHTIPPPNGMNLKNNHSESSNRSYRDSGIPDSDSHS 649

RESULT 10

AAB55989
 ID AAB55989 standard; Protein; 649 AA.

AC AAB55989;

DT 08-MAR-2001 (first entry)

XX Skin cell protein, SEQ ID NO: 305.

XX Mouse; skin cell; cytostatic; antiinflammatory; anti-HIV;
 KW nontropic; neuroprotective; vulnery; immunomodulatory; vaccine;
 KW keratinocyte growth stimulation; cancer; angiogenesis inhibition;
 KW inflammation; neurological disease.

XX Mus sp.

XX WO200069884-A2.

XX 23-NOV-2000.

XX 15-MAY-2000; 2000WO-NZ00075.

XX 14-MAY-1999; 99US-0312283.

XX (GENE-) GENESIS RES & DEV CORP LTD.

XX Watson JD, Strachan L, Onrust R, Sleeman M, Kumble KD, Murison JG;

XX WPI; 2001-007495/01.

XX N-PSDB; AAC99688.

XX New isolated polynucleotide used in the identification of genetic
 PT disorders and encoding polypeptides used for treating inflammatory
 PT disease, cancer and neurological diseases -

XX Claim 4; Page 246-247; 352pp; English.

XX The present sequence is a polypeptide which is expressed in
 CC mammalian skin cells. The polypeptide is useful for stimulating

CC keratinocyte growth and motility, inhibiting the growth of cancer cells,
 CC modulating angiogenesis, inhibiting angiogenesis and vascularization of
 CC tumours, modulating skin inflammation, stimulating the growth of
 CC epithelial cells, inhibiting the binding of human immunodeficiency virus
 CC (HIV)-1 to leukocytes, and treating inflammatory disease, cancer and
 CC neurological diseases. The polynucleotide can be used as a marker, in
 CC the identification of genetic disorders, and for the design of
 CC oligonucleotides for examining expression patterns.

XX Sequence 649 AA;

Query Match 96.6%; Score 3284; DB 22; Length 649;

Best Local Similarity 96.5%; Pred. No. 1.2e-279;

Matches 626; Conservative 11; Mismatches 12; Indels 0; Gaps 0;

QY 1 MISAWSEFLGTGKIGLFQVAPLSVWAKSCPSVCRCAGFIYCNDRLTSTPTGIPEDA 60

Db 1 MISPAWSFLGTGKIGLFQVAPLSVWAKSCPSVCRCAGFIYCNDRLTSTPTGIPEDA 60

QY 61 TTYLQNNCINNAGIPSDLNKLNKVERIYIYHNSLDEFTNLPKYVKELHLOENNRITIT 120

Db 61 TTYLQNNCINNAGIPSDLNKLNKVERIYIYHNSLDEFTNLPKYVKELHLOENNRITIT 120

QY 121 YDSLSKIPYLEELHLDNNSVSAVSIIEGAFRDSNYLRLFLSRNHLSTIPWGLPTIEEL 180

Db 121 YDSLSKIPYLEELHLDNNSVSAVSIIEGAFRDSNYLRLFLSRNHLSTIPWGLPTIEEL 180

QY 181 RLDDNRISTISSPSLOGLTSLKRLVLDGNLNNHGLGDKVFNFLVNLTELSELVRSNLTA 240

Db 181 RLDDNRISTISSPSLOGLTSLKRLVLDGNLNNHGLGDKVFNFLVNLTELSELVRSNLTA 240

QY 241 PVNLPOTNLRLKLYLDQNHINRVPNFAFSLRYLYRLDMSNNLSNLPQGFDDLDNITQL 300

Db 241 PVNLPOTNLRLKLYLDQNHINRVPNFAFSLRYLYRLDMSNNLSNLPQGFDDLDNITQL 300

QY 301 ILRNNPWYCGCKMKWVRDLQSLPVKNVVRGLMCOAPEKVRGMAIKDLNAELFDCKDSGI 360

Db 301 ILRNNPWYCGCKMKWVRDLQSLPVKNVVRGLMCOAPEKVRGMAIKDLNAELFDCKDSGI 360

QY 361 VSTIQTITPAINTVYPAQGWPAVTKQPDIKNPKLTQDQTTGSPSRKTIITVKSPTS 420

Db 361 VSTIQTITPAINTVYPAQGWPAVTKQPDIKNPKLTQDQTTGSPSRKTIITVKSPTS 420

QY 421 DTIHISWALPMTALRLSLWKLGHSPAFGSIETIVTGERSEYIVTALPESPVKCV 480

Db 421 DTIHISWALPMTALRLSLWKLGHSPAFGSIETIVTGERSEYIVTALPESPVKCV 480

QY 481 PMETSNLYLDFETPVCIEETAPLRMTNPTTLNRQEKPEYKKNPNPLAAIIGGVALV 540

Db 481 PMETSNLYLDFETPVCIEETAPLRMTNPTTLNRQEKPEYKKNPNPLAAIIGGVALV 540

QY 541 TIALALVGVYHRNGSLFSRNCAYSKGRRKDDYAEAGTKKDNSILEIRETSFQMLPIS 600

Db 541 TIALALVGVYHRNGSLFSRNCAYSKGRRKDDYAEAGTKKDNSILEIRETSFQMLPIS 600

QY 601 NEPISKEEFVIHTIPPPNGMNLKNNHSESSNRSYRDSGIPDSDSHS 649

Db 601 NEPISKEEFVIHTIPPPNGMNLKNNHSESSNRSYRDSGIPDSDSHS 649

RESULT 11

AAB72189
 ID AAB72189 standard; Protein; 649 AA.

AC AAB72189;

XX 04-APR-2002 (first entry)

XX Murine protein isolated from skin cells SEQ ID NO: 305.

XX Human; rat; mouse; skin cell; skin wound; cancer; growth defect;
 KW developmental defect; inflammatory disease; dermatological; vulnery;
 KW immunomodulator; anti-inflammatory; cytostatic; neuroprotective.

XX OS Mus sp.
 XX PN WC200190357-Al.
 XX PD 29-NOV-2001.
 XX PF 24-MAY-2001; 2001WO-NZ00099.
 XX PR 24-MAY-2000; 2000US-20650P.
 XX PR 25-JUL-2000; 2000US-221232P.
 XX PA (GENE-) GENESIS RES & DEV CORP LTD.
 XX PI Watson JD, Strachan L, Sleeman M, Orrust R, Murison JG, Kumble KD;
 XX DR WPI; 2002-122020/16.
 XX PT New polynucleotides and polypeptides encoded by the polynucleotides
 PT isolated from skin cells, useful for treating skin wounds, cancers,
 PT growth and developmental defects, inflammatory diseases, or for
 PT modulating immune responses
 XX PS Example 2; Page 202-203; 466pp; English.
 XX SS The present invention provides the protein and coding sequences of cDNAs
 CC isolated from human, murine and rat skin cell libraries. The sequences
 CC can be used in the development of therapeutic agents useful in the
 CC treatment of skin diseases, including skin wounds, cancer, growth
 CC defects, developmental defects and inflammatory diseases. The proteins
 CC have important roles in the induction of hair growth, cell proliferation
 CC and cell-cell interaction, in maintaining tissue integrity, in wound
 CC healing and in modulating immune responses. The present sequence is a
 CC polypeptide of the invention.
 XX SQ Sequence 649 AA;

Query Match 96.6%; Score 3284; DB 23; Length 649;
 Best Local Similarity 96.5%; Pred. NO. 1.2e-279;
 Matches 626; Conservative 11; Mismatches 12; Indels 0; Gaps 0;

QY 1 MISAWSTFLGTGTLGFLQVAPLSVMAKSCPSVCRDAGFIYCNDRFLTSIPTGIPEDA 60
 DB 1 MISPAWSTFLGTGTLGFLQVAPLSVMAKSCPSVCRDAGFIYCNDRFLTSIPTGIPEDA 60
 QY 61 TTLYLQNNQINNAGIPSDLNKLVKRIYLYHNSLDEPTNLPKYKELHQLQENIRIT 120
 DB 61 TTLYLQNNQINNAGIPSDLNKLVKRIYLYHNSLDEPTNLPKYKELHQLQENIRIT 120
 QY 121 YDLSKIPYLELHLDNNSVSAVSBEGAFRDSNYLRLFLSRNHLSTIPWGLPRTIEEL 180
 DB 121 YDLSKIPYLELHLDNNSVSAVSBEGAFRDSNYLRLFLSRNHLSTIPWGLPRTIEEL 180
 QY 181 RLDDNRISTISSPSTQGTSLKRLVLDGNLNNHGLGDKVFNVLNLTLSLRNSLRAA 240
 DB 181 RLDDNRISTISSPSTQGTSLKRLVLDGNLNNHGLGDKVFNVLNLTLSLRNSLRAA 240
 QY 241 PVNLPTNLRLKLYLQDNIHNRFPNFAFSLRYLRIDMSNNLSNLPQGFDDLDNIQL 300
 DB 241 PVNLPTNLRLKLYLQDNIHNRFPNFAFSLRYLRIDMSNNLSNLPQGFDDLDNIQL 300
 QY 301 ILRNNPWCCKKWKWRWQSLQSPVKVNVRLGMLCOAPEKVRGMAIKDLNAELFDCKDSGI 360
 DB 301 ILRNNPWCCKKWKWRWQSLQSPVKVNVRLGMLCOAPEKVRGMAIKDLNAELFDCKDSGI 360
 QY 361 VSTIQLTTPAIPVNYPAQGWPAVVKQPDIKNPKLTQDQQTGSPSKTITITVKSQVTS 420
 DB 361 VSTIQLTTPAIPVNYPAQGWPAVVKQPDIKNPKLTQDQQTGSPSKTITITVKSQVTS 420
 QY 421 DTIHISWKLALPMTALRLSLWKLGHSPAFSGSITEITIVTGERSEYLVTALEPDSPKVCMV 480
 DB 421 DTIHISWKLALPMTALRLSLWKLGHSPAFSGSITEITIVTGERSEYLVTALEPDSPKVCMV 480

QY 481 PMETSNLYLDFETPVCIETETAPLRMYNPTTTLNREQEKEPYKPNPLPLAALIGGVALV 540
 DB 481 PMETSNLYLDFETPVCIETETAPLRMYNPTTTLNREQEKEPYKPNPLPLAALIGGVALV 540
 QY 541 TIALALVCMYVHRNGSLFSRNCAYSKGRRKKDDYABAGTKKDNSILEIRETSFQMLPIS 600
 DB 541 TIALALVCMYVHRNGSLFSRNCAYSKGRRKKDDYABAGTKKDNSILEIRETSFQMLPIS 600
 QY 601 NEPIKSEEFVHTIFFPNGNLYKNNHSESSNSRYSRDSGIPDSHSHS 649
 DB 601 NEPIKSEEFVHTIFFPNGNLYKNNHSESSNSRYSRDSGIPDSHSHS 649

RESULT 12
 AAU19626
 ID AAU19626 standard; Protein; 623 AA.
 XX AC AAU19626;
 XX DT 04-DEC-2001 (first entry)
 XX DE Human novel extracellular matrix protein, Seq ID No 276.
 XX KW Human; secreted extracellular matrix protein; immunomodulatory;
 KW Anti-HIV; antianemic; antirheumatic; antisclerotic; cardiac; vascular;
 KW cerebroprotective; thrombolytic; antimicrobial; ophthalmic; cytostatic;
 KW antiallergics; immune/autoimmune disease; HIV infection; anaemia;
 KW human immunodeficiency virus; rheumatoid arthritis; multiple sclerosis;
 KW cancer; hyperproliferative disorder; breast neoplasia; melanoma;
 KW Sezary syndrome; Gaucher's disease; neurological diseases;
 KW Alzheimer's disease; Parkinson's disease; cardiovascular disorder;
 KW cardiac arrest; tachycardia; angina; infection; corneal infections;
 KW wound healing; immunogen; gene therapy; antisense; food additive.
 XX OS Homo sapiens.
 XX PN WC200155368-Al.
 XX PD 02-AUG-2001.
 XX PF 17-JAN-2001; 2001WO-US01348.
 XX PR 31-JAN-2000; 2000US-0179065.
 PR 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184664.
 PR 02-MAR-2000; 2000US-0186350.
 PR 16-MAR-2000; 2000US-0189874.
 PR 17-MAR-2000; 2000US-0190076.
 PR 18-APR-2000; 2000US-0198123.
 PR 19-MAY-2000; 2000US-0205515.
 PR 07-JUN-2000; 2000US-0209467.
 PR 28-JUN-2000; 2000US-0214886.
 PR 30-JUN-2000; 2000US-0215135.
 PR 07-JUL-2000; 2000US-0216647.
 PR 11-JUL-2000; 2000US-0217487.
 PR 11-JUL-2000; 2000US-0217496.
 PR 14-JUL-2000; 2000US-0218290.
 PR 26-JUL-2000; 2000US-0220963.
 PR 26-JUL-2000; 2000US-0220964.
 PR 14-AUG-2000; 2000US-0224518.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0225213.
 PR 14-AUG-2000; 2000US-0225214.
 PR 14-AUG-2000; 2000US-0225266.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225447.
 PR 14-AUG-2000; 2000US-0225757.
 PR 14-AUG-2000; 2000US-0225758.
 PR 14-AUG-2000; 2000US-0225759.
 PR 18-AUG-2000; 2000US-0226279.

PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226686.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246509.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.

PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Barash SC, Ruben SM;
XX WPL: 2001-465572/50.
XX N-PSDB: AAS31197.
PT Nucleic acid molecules encoding human secreted extracellular matrix
PT proteins, used in preventing, treating or ameliorating a disorder, e.g.
PT Alzheimer's and Parkinson's diseases and cancers -
PS Claim 11; SEQ ID No 276; 577pp; English.
XX The invention relates to isolated nucleic acid molecules encoding
XX novel human secreted extracellular matrix proteins (SPs). The
XX polynucleotides and proteins are used to prevent, treat a medical
XX condition in e.g. humans, mice, rabbits, goats, horses, cats, dogs,
XX chickens or sheep. For example, disorders associated with decreased
XX expression of SPs. The SP polynucleotide or a vector expressing them may
XX be administered to treat diseases by gene therapy. Antisense molecules
XX may be administered to down regulate expression of SPs by binding with
XX the cells own genes and preventing their expression. The polynucleotides
XX may also be used as DNA probes in diagnostic assays. The SPs may also be
XX used as antigens to produce antibodies and to identify modulators
XX (agonists and antagonists) of the SPs. The anti-(SP) antibodies and
XX antagonists may also be used to down regulate expression and activity of
XX SP and as diagnostic agents for detecting the presence of SPs in samples.
XX The disorders include for example: immune/autoimmune diseases (e.g. HIV
XX (human immunodeficiency virus) infections, anaemia, rheumatoid arthritis
XX and multiple sclerosis), cancers and hyperproliferative disorders (e.g.
XX melanomas, neoplasms of the breast or liver, Sezary syndrome and
XX Gaucher's disease), neurological diseases (e.g. Alzheimer's disease,
XX Parkinson's disease) cardio-/cerebrovascular disorders (e.g. cardiac
XX arrest, tachycardia and angina), infections caused by bacteria, viruses
XX and fungi and ocular disorders (e.g. corneal infections). Other uses
XX include wound healing, maintenance of organs before transplantation,
Query Match 96.3%; Score 3274; DB 22; Length 623;
Best Local Similarity 100.0%; Pred. No. 8.5e-279;
Matches 623; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 27 MAKSPSVCRDAGFYCNDRELTSPGIPEDATTIYLQNNQINNAGIPSDLNKLKVE 86
Db 1 MAKSPSVCRDAGFYCNDRELTSPGIPEDATTIYLQNNQINNAGIPSDLNKLKVE 60
QY 87 RYIYHNSLDEFTNPKYVKELHLOENNIRTIYDLSKIPYBELHLDNDSVAYSIE 146
Db 61 RYIYHNSLDEFTNPKYVKELHLOENNIRTIYDLSKIPYBELHLDNDSVAYSIE 120
QY 147 EGAFRDSNYRLFLGRNHLSTIPWGLPRTTEELRLDDNKRISTSPSLOGLTSKLKVL 206
Db 121 EGAFRDSNYRLFLGRNHLSTIPWGLPRTTEELRLDDNKRISTSPSLOGLTSKLKVL 180
QY 207 DGNLNNHGLGDKVFFNLVNLVELSILVRNSLTAAPVNLPGTNLRKLYLQDNHINRPPNA 266
Db 181 DGNLNNHGLGDKVFFNLVNLVELSILVRNSLTAAPVNLPGTNLRKLYLQDNHINRPPNA 240
QY 267 FSYLRQLYRLDMNNLSNLPQIGIFDDLDNITQLILRNPNWTCGCKMKWRDQLQSLPVK 326
Db 241 FSYLRQLYRLDMNNLSNLPQIGIFDDLDNITQLILRNPNWTCGCKMKWRDQLQSLPVK 300
QY 327 VNVRLGMCQPEKVRGMAIKDLNLELFDCKDSGIVSTIQTITTAIPNTVYPAGOGWPAPVT 386
Db 301 VNVRLGMCQPEKVRGMAIKDLNLELFDCKDSGIVSTIQTITTAIPNTVYPAGOGWPAPVT 360
QY 387 KOPDIKNPKLTQDOQTGSPSRKTTITVKSVDSTIHSKWLALPMTALRLSLKLGHS 446
Db 361 KOPDIKNPKLTQDOQTGSPSRKTTITVKSVDSTIHSKWLALPMTALRLSLKLGHS 420
QY 447 PAFGSITETIVTGERSEYLVTALEPDSYKVCNMPAETSNIYLFDETPYCETETAPIRM 506
Db 421 PAFGSITETIVTGERSEYLVTALEPDSYKVCNMPAETSNIYLFDETPYCETETAPIRM 480
QY 507 YNPFTTLNRQEKPEKYNPLALAIIGGAVALTIALALVCVYVHRNGSLFERNCAYS 566
Db 481 YNPFTTLNRQEKPEKYNPLALAIIGGAVALTIALALVCVYVHRNGSLFERNCAYS 540
QY 567 KGRKKDDYAPAGTKKDNLSLEIRETSQMLPISNEPISKEEFVHTIPPNGMNLKYN 626
Db 541 KGRKKDDYAPAGTKKDNLSLEIRETSQMLPISNEPISKEEFVHTIPPNGMNLKYN 600
QY 627 HSESSNKSRYDSCGIPDSHSHS 649
Db 601 HSESSNKSRYDSCGIPDSHSHS 623
RESULT 13
ID ABP47846
XX ABP47846 standard; Protein; 623 AA.
AC ABP47846;
XX
XX 23-AUG-2002 (first entry)
XX Human polypeptide SEQ ID NO 276.
XX
KW Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;
KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnary;
KW antiparkinsonian; antisticking; antianseptic; antiarthritic; cancer;
KW antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;
KW antiallergic; antidiabetic; antilucer; anticonvulsant; antifungal;
KW antiparasitic; cardiant; immune disorder; cardiovascular disorder;
KW neurological disease; infection; nephrotropic; gene therapy; vaccine.
XX
XX Homo sapiens.
XX
XX US2002042386-A1.
XX
XX 11-APR-2002.
XX
XX 17-JAN-2001; 2001US-0764870.
XX
XX 31-JAN-2000; 2000US-179065P.
PR

PR 04-FEB-2000; 2000US-180628P.
PR 28-JUN-2000; 2000US-214886P.
PR 07-JUL-2000; 2000US-216647P.
PR 07-JUL-2000; 2000US-216880P.
PR 11-JUL-2000; 2000US-217487P.
PR 11-JUL-2000; 2000US-217496P.
PR 14-JUL-2000; 2000US-218290P.
PR 26-JUL-2000; 2000US-220963P.
PR 26-JUL-2000; 2000US-220964P.
PR 14-AUG-2000; 2000US-224518P.
PR 14-AUG-2000; 2000US-224519P.
PR 14-AUG-2000; 2000US-225267P.
PR 14-AUG-2000; 2000US-225268P.
PR 14-AUG-2000; 2000US-225270P.
PR 14-AUG-2000; 2000US-225447P.
PR 14-AUG-2000; 2000US-225757P.
PR 14-AUG-2000; 2000US-225758P.
PR 22-AUG-2000; 2000US-226868P.
PR 30-AUG-2000; 2000US-228924P.
PR 01-SEP-2000; 2000US-229287P.
PR 01-SEP-2000; 2000US-229343P.
PR 01-SEP-2000; 2000US-229344P.
PR 01-SEP-2000; 2000US-229345P.
PR 05-SEP-2000; 2000US-229509P.
PR 05-SEP-2000; 2000US-229513P.
PR 08-SEP-2000; 2000US-231413P.
PR 21-SEP-2000; 2000US-234223P.
PR 21-SEP-2000; 2000US-234274P.
PR 25-SEP-2000; 2000US-234997P.
PR 27-SEP-2000; 2000US-235834P.
PR 29-SEP-2000; 2000US-236327P.
PR 29-SEP-2000; 2000US-236367P.
PR 29-SEP-2000; 2000US-236368P.
PR 29-SEP-2000; 2000US-236369P.
PR 29-SEP-2000; 2000US-236370P.
PR 02-OCT-2000; 2000US-236802P.
PR 02-OCT-2000; 2000US-237037P.
PR 02-OCT-2000; 2000US-237038P.
PR 02-OCT-2000; 2000US-237039P.
PR 02-OCT-2000; 2000US-237040P.
PR 13-OCT-2000; 2000US-239935P.
PR 20-OCT-2000; 2000US-240960P.
PR 20-OCT-2000; 2000US-241785P.
PR 20-OCT-2000; 2000US-241809P.
PR 01-NOV-2000; 2000US-244617P.
PR 17-NOV-2000; 2000US-249299P.
PR 08-DEC-2000; 2000US-251856P.
PR 08-DEC-2000; 2000US-251868P.
PR 08-DEC-2000; 2000US-251869P.
XX
PA (ROSE/) ROSEN C A.
PA (RUBE/) RUBEN S M.
PA (BARA/) BARASH S C.
PI Rosen CA, Ruben SM, Barash SC;
XX
XX WPI; 2002-470713/50.
DR N-PSDB; ABQ66521.
XX
PT New nucleic acid encoding human proteins, useful for diagnosis,
PT treatment and prevention of e.g. osteoporosis, also related
XX polypeptides and antibodies
XX
PS Claim 11; SEQ ID NO 276; 235pp + Sequence Listing; English.
XX
CC The invention relates to novel genes (ABQ66521-ABQ66785) and proteins
CC (ABP47846-ABP48110) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful
CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;

CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;
CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
CC and parasitic infections.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from USPTO at seqdata.uspto.gov/sequence.html?DocID=99909764870.
XX
SQ Sequence 623 AA;

Query Match 96.3%; Score 3274; DB 23; Length 623;
Best Local Similarity 100.0%; Pred. No. 8.5e-279;
Matches 623; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 27 MAKSCPSVCRDAGFYCNDREFTSPTGIPEDATLYLQNNQINNAGIPSDKLLKLYE 86
Db 1 MAKSCPSVCRDAGFYCNDREFTSPTGIPEDATLYLQNNQINNAGIPSDKLLKLYE 60
QY 87 RYLYHNSLDEFPNPKYVKELHQLQNNIRITVDSLSKIPYLEELHDDNSVSAVSIE 146
Db 61 RYLYHNSLDEFPNPKYVKELHQLQNNIRITVDSLSKIPYLEELHDDNSVSAVSIE 120
QY 147 EGAFRDSNYRLRLFLSRNHLSTIPWGLPRTIEELRDDNRISTISSPSLQGLTSKRLVL 206
Db 121 EGAFRDSNYRLRLFLSRNHLSTIPWGLPRTIEELRDDNRISTISSPSLQGLTSKRLVL 180
QY 207 DGNLLNNGHGLDKYFFNLVNLTELSVRSNLSLAAPVNLPGTNLRKLYLQDNHINRVPNA 266
Db 181 DGNLLNNGHGLDKYFFNLVNLTELSVRSNLSLAAPVNLPGTNLRKLYLQDNHINRVPNA 240
QY 267 FSYLRQLYRLDMSNNLSNLPQIGIPDDLNITQLLRNPNWYCGCKMKVRWLOSLPVK 326
Db 241 FSYLRQLYRLDMSNNLSNLPQIGIPDDLNITQLLRNPNWYCGCKMKVRWLOSLPVK 300
QY 327 VNVRLGMCQAEKPYRGMAIKDLNAELFDCDGSIVSTIITTAIPNTVYPAQGWPAVPT 386
Db 301 VNVRLGMCQAEKPYRGMAIKDLNAELFDCDGSIVSTIITTAIPNTVYPAQGWPAVPT 360
QY 387 KOPDKNPKLKDOQTGSPSRKTIITVKSVTSDTHISWKLALPMTALRLSWLKLGH 446
Db 361 KOPDKNPKLKDOQTGSPSRKTIITVKSVTSDTHISWKLALPMTALRLSWLKLGH 420
QY 447 PAFGSIETIIVTGERSEYLVTALEPDSYKVCWMPMETSNIYLFDEPVCIEFETAPLRM 506
Db 421 PAFGSIETIIVTGERSEYLVTALEPDSYKVCWMPMETSNIYLFDEPVCIEFETAPLRM 480
QY 507 YNPFTTLNREQEKPEYKPNPLPLAAIIGGAVALTIALALVCWYVHRNGSLFSRNCAYS 566
Db 481 YNPFTTLNREQEKPEYKPNPLPLAAIIGGAVALTIALALVCWYVHRNGSLFSRNCAYS 540
QY 567 KGRKKDDYAGAGTKKNSILETRETFSOMLPISNEPISKEEFVIHTIPPPNGMNLKYN 626
Db 541 KGRKKDDYAGAGTKKNSILETRETFSOMLPISNEPISKEEFVIHTIPPPNGMNLKYN 600
QY 627 HSESSSNRSYRDSGIPDSHSHS 649
Db 601 HSESSSNRSYRDSGIPDSHSHS 623

RESULT 14
AAB93199
ID AAB93199 standard; Protein; 623 AA.
XX
AC AAB93199;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human protein sequence SEQ ID NO:12162.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.

XX Homo sapiens.
OS EF1074617-A2.
PN
XX
PD 07-FEB-2001.
XX
XX 28-JUL-2000; 2000EP-0116126.
XX
XX 29-JUL-1999; 99JP-0248036.
PR 27-AUG-1999; 99JP-0300253.
PR 11-JAN-2000; 2000JP-0118776.
PR 02-MAY-2000; 2000JP-0183767.
PR 09-JUN-2000; 2000JP-0241899.
XX
XX (HELI-) HELIX RES INST.
XX
XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX WPI; 2001-318749/34.
XX
XX Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs -
XX
PS Claim 8; SEQ ID 12162; 2537pp + CD ROM; English.
XX
XX The present invention describes primer sets for synthesizing 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dr primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination
CC of the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.
XX
SQ Sequence 623 AA;

Query Match 96.1%; Score 3269; DB 22; Length 623;
Best Local Similarity 99.8%; Pred. No. 2.3e-278;
Matches 622; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 27 MAKSCPSVCRDAGFYCNDREFTSPTGIPEDATLYLQNNQINNAGIPSDKLLKLYE 86
Db 1 MAKSCPSVCRDAGFYCNDREFTSPTGIPEDATLYLQNNQINNAGIPSDKLLKLYE 60
QY 87 RYLYHNSLDEFPNPKYVKELHQLQNNIRITVDSLSKIPYLEELHDDNSVSAVSIE 146
Db 61 RYLYHNSLDEFPNPKYVKELHQLQNNIRITVDSLSKIPYLEELHDDNSVSAVSIE 120
QY 147 EGAFRDSNYRLRLFLSRNHLSTIPWGLPRTIEELRDDNRISTISSPSLQGLTSKRLVL 206
Db 121 EGAFRDSNYRLRLFLSRNHLSTIPWGLPRTIEELRDDNRISTISSPSLQGLTSKRLVL 180
QY 207 DGNLLNNGHGLDKYFFNLVNLTELSVRSNLSLAAPVNLPGTNLRKLYLQDNHINRVPNA 266
Db 181 DGNLLNNGHGLDKYFFNLVNLTELSVRSNLSLAAPVNLPGTNLRKLYLQDNHINRVPNA 240

QY 267 PSYLRQLYRLDMSNNLSNLPQGIFFDDLDNITQILRLNNPWYCGCKMKWVRDLQSLFVK 326
DB 241 PSYLRQLYRLDMSNNLSNLPQGIFFDDLDNITQILRLNNPWYCGCKMKWVRDLQSLFVK 300
QY 327 VNVRLGMCQAEKVRGMAIKDLNAELFCKDSGIVSTIQTITAPNTVTPAQGWPAFVT 386
DB 301 VNVRLGMCQAEKVRGMAIKDLNAELFCKDSGIVSTIQTITAPNTVTPAQGWPAFVT 360
QY 387 KOPDLKPKLKDOOTGSPSRKTITIVKSVTSDTHISWKLALPMTALRLSWLKLGH 446
DB 361 KOPDLKPKLKDOOTGSPSRKTITIVKSVTSDTHISWKLALPMTALRLSWLKLGH 420
QY 447 PAFGSITETIVTGRSEYLVTALEPSPYKVCMPMETSNIYLFDETPVCIEETAPLRM 506
DB 421 PAFGSITETIVTGRSEYLVTALEPSPYKVCMPMETSNIYLFDETPVCIEETAPLRM 480
QY 507 YNPFTTLNREQEKPEPKYKPNPLPLAALIGGAVALTITALLALVCWYVHRNGSLFERN 566
DB 481 YNPFTTLNREQEKPEPKYKPNPLPLAALIGGAVALTITALLALVCWYVHRNGSLFERN 540
QY 567 KGRKKDYAAGYKDKNSILETRETSTFOMLPISNEPISKEEFVHTIETPPNGMNLKNN 626
DB 541 KGRKKDYAAGYKDKNSILETRETSTFOMLPISNEPISKEEFVHTIETPPNGMNLKNN 600
QY 627 HSESSNRSYRDSIGIPDSHSHS 649
DB 601 HSESSNRSYRDSIGIPDSHSHS 623

RESULT 15
AAM93814
ID AAM93814 standard; Protein; 627 AA.
AC AAM93814;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human polypeptide, SEQ ID NO: 3863.
XX
KW Human; full length cDNA; cDNA synthesis; oligo-capping.
XX
OS Homo sapiens.
XX
PN BFL130094-A2.
XX
PD 05-SEP-2001.
XX
PF 07-JUL-2000; 2000EP-0114089.
XX
PR 08-JUL-1999; 99JP-0194486.
PR 11-JAN-2000; 2000JP-0118774.
PR 02-MAY-2000; 2000JP-0183765.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX
WPI: 2001-524255/58.
N-PSDB; AAK94770.

830 Primers useful for synthesizing full length cDNA clones and their use in genetic manipulation -
XX
PS Claim 8; SEQ ID NO 3863; 1380pp + sequence listing; English.
XX
CC The invention relates to primers for synthesizing full length cDNA clones. 830 cDNA molecules encoding a human protein have been isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have been determined. Primers for synthesizing the full length cDNA are useful for clarifying the function of the protein encoded by the cDNA. The full length clones were obtained by construction of full

CC length enriched cDNA libraries that were synthesised by the oligo-capping method. The primers enable the production of the full length cDNA easily without any special methods. The present sequence is a polypeptide encoded by a full length human cDNA of the invention.
CC Note: The sequence data for this patent did not form part of the printed CC specification, but was obtained in CD-ROM format directly from EPO.
XX
SQ Sequence 627 AA;

Query Match 95.9%; Score 3261; DB 22; Length 627;
Best Local Similarity 99.5%; Pred. No. 1.2e-277;
Matches 623; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 1 MISAAWSIFLIGTKIGLFLQVAPLSVMAKSCPSVCRDAGFIYCNDRFLTSTIPGIPEDA 60
DB 1 MISAAWSIFLIGTKIGLFLQVAPLSVMAKSCPSVCRDAGFIYCNDRFLTSTIPGIPEDA 60
QY 61 TTYLQNNQINNAGIPSDKLLKVERIYLYHNSLDEFPNLPKYVKELHQLQENNIRIT 120
DB 61 TTYLQNNQINNAGIPSDKLLKVERIYLYHNSLDEFPNLPKYVKELHQLQENNIRIT 120
QY 121 YDLSKIPYLEELHLDNSVSAVSTEEGAFRDSNYLRLFLSRNHLSTIPWGLPTIEEL 180
DB 121 YDLSKIPYLEELHLDNSVSAVSTEEGAFRDSNYLRLFLSRNHLSTIPWGLPTIEEL 180
QY 181 RLDDNRISTISSPSLQGLTSLKRLVLDGNLNNHGLGDKVFFNLVNLTELSTLVRNSLTAA 240
DB 181 RLDDNRISTISSPSLQGLTSLKRLVLDGNLNNHGLGDKVFFNLVNLTELSTLVRNSLTAA 240
QY 241 PYNLPGTNLRKLILODNHINRVPPNAPSYLRQLYRLDMSNNLSNLPQGIFFDDLDNITQL 300
DB 241 PYNLPGTNLRKLILODNHINRVPPNAPSYLRQLYRLDMSNNLSNLPQGIFFDDLDNITQL 300
QY 301 ILRNNPWYCGCKMKWVRDLQSLPYKVNVRGLMCQAEKVRGMAIKDLNAELFCKDSGI 360
DB 301 ILRNNPWYCGCKMKWVRDLQSLPYKVNVRGLMCQAEKVRGMAIKDLNAELFCKDSGI 360
QY 361 VSTIQTITAPNTVTPAQGWPAFVTQPDIKNPKLTKDQQTGSPSRKTITITVKSVT 420
DB 361 VSTIQTITAPNTVTPAQGWPAFVTQPDIKNPKLTKDQQTGSPSRKTITITVKSVT 420
QY 421 DTIHISWKLALPMTALRLSWLKLGHSPAFGSTITETIVTGRSEYLVTALEPSPYKVCW 480
DB 421 DTIHISWKLALPMTALRLSWLKLGHSPAFGSTITETIVTGRSEYLVTALEPSPYKVCW 480
QY 481 PMETSNIYLFDETPVCIEETAPLRMYNFTTLNREQEKPEPKYKPNPLPLAALIGGAV 540
DB 481 PMETSNIYLFDETPVCIEETAPLRMYNFTTLNREQEKPEPKYKPNPLPLAALIGGAV 540
QY 541 TIALALVCWYVHRNGSLFERNKAYSGKRRKDDYABAGYKDKNSILETRETSTFOMLPIS 600
DB 541 TIALALVCWYVHRNGSLFERNKAYSGKRRKDDYABAGYKDKNSILETRETSTFOMLPIS 600
QY 601 NEPISKEEFVHTIETPPNGMNLKNN 626
DB 601 NEPISKEEFVHTIETPPNGMNLKNN 626

Search completed: April 23, 2003, 11:05:16
Job time : 42 secs

GenCore version 5.1.4 p5 4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 11:04:36 ; Search time 17 seconds
(without alignments)
1123.263 Million cell updates/sec

Title: US-10-004-551-6
Perfect score: 3400
Sequence: 1 MISAWSIFLIGTKIGLFQ.....SSNSRYSRDSGIPSDSHS 649

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA.*
1: /cgn2_6/prodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/prodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/prodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/prodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/prodata/1/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/prodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	3284	96.6	649	US-09-188-930-305	Sequence 305, App
2	346	16.1	110	US-09-188-930-124	Sequence 124, App
3	397	11.7	708	US-09-131-648-2	Sequence 2, Appl
4	377.5	11.1	1523	US-09-182-024A-2	Sequence 2, Appl
5	354.5	10.4	673	US-09-063-950-2	Sequence 2, Appl
6	345.5	10.2	1480	US-09-182-024A-5	Sequence 5, Appl
7	344.5	10.1	1480	US-09-191-647-7	Sequence 7, Appl
8	344.5	10.1	1480	US-09-540-245A-7	Sequence 7, Appl
9	344.5	10.1	1480	US-09-540-153-7	Sequence 7, Appl
10	344.5	10.1	1480	PCT-US91-09055-2	Sequence 2, Appl
11	343.5	10.1	1525	US-09-191-647-2	Sequence 2, Appl
12	343.5	10.1	1525	US-09-540-245A-2	Sequence 2, Appl
13	343.5	10.1	1525	US-09-540-153-2	Sequence 2, Appl
14	342	10.1	282	US-08-442-063A-45	Sequence 45, Appl
15	342	10.1	307	US-08-442-063A-48	Sequence 48, Appl
16	342	10.1	333	US-08-442-063A-27	Sequence 27, Appl
17	342	10.1	342	US-08-272-919-2	Sequence 2, Appl
18	342	10.1	342	US-08-619-916-2	Sequence 2, Appl
19	342	10.1	342	PCT-US95-08542-2	Sequence 2, Appl
20	342	10.1	359	US-08-303-238-4	Sequence 4, Appl
21	342	10.1	359	US-08-458-834-4	Sequence 4, Appl
22	335.5	9.9	376	US-08-303-238-1	Sequence 1, Appl
23	335.5	9.9	376	US-08-458-834-1	Sequence 1, Appl
24	333.5	9.8	375	US-08-303-238-2	Sequence 2, Appl
25	333.5	9.8	375	US-08-458-834-2	Sequence 2, Appl
26	320.5	9.4	373	US-09-724-864-43	Sequence 43, Appl
27	316.5	9.3	1091	US-08-986-485-5	Sequence 5, Appl

28	314.5	9.2	605	1	US-08-190-802A-49	Sequence 49, Appl
29	314.5	9.2	605	4	US-08-477-346-49	Sequence 49, Appl
30	314.5	9.2	605	4	US-08-477-346-49	Sequence 49, Appl
31	314.5	9.2	605	4	US-08-477-346-49	Sequence 49, Appl
32	311.5	9.2	605	4	US-09-063-950-5	Sequence 5, Appl
33	306	9.0	603	1	US-08-190-802A-50	Sequence 50, Appl
34	306	9.0	603	4	US-08-477-346-50	Sequence 50, Appl
35	306	9.0	603	4	US-08-477-346-50	Sequence 50, Appl
36	306	9.0	603	4	US-08-477-346-50	Sequence 50, Appl
37	303	8.9	603	6	US-08-487-072A-50	Sequence 50, Appl
38	299.5	8.8	236	1	US-08-442-063A-42	Sequence 42, Appl
39	294	8.6	1101	3	US-08-986-485-2	Sequence 2, Appl
40	283.5	8.3	368	1	US-08-303-238-3	Sequence 3, Appl
41	283.5	8.3	368	4	US-08-458-834-3	Sequence 3, Appl
42	277.5	8.2	368	6	US-08-458-834-3	Sequence 3, Appl
43	268.5	7.9	224	5	PCT-US91-09055-4	Sequence 4, Appl
44	260	7.6	231	3	US-08-986-485-7	Sequence 7, Appl
45	259.5	7.6	222	5	PCT-US91-09055-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1
US-09-188-930-305
; Sequence 305, Application US/09188930A
; Patent No. 6150502

GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James Greg
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; TITLE OF INVENTION: and Methods For Their Use
; FILE REFERENCE: 11000.1011c1
; CURRENT APPLICATION NUMBER: US/09/188.930A
; CURRENT FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 348
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 305
; LENGTH: 649
; TYPE: PRT
; ORGANISM: Mouse
US-09-188-930-305

Query Match	96.6%	Score 3284;	DB 4;	Length 649;
Best Local Similarity	96.5%	Pred. No. 1.4e-287;		
Matches 626;	Conservative 11;	Mismatches 12;	Indels 0;	Gaps 0;
QY	1	MISAWSIFLIGTKIGLFQVAPLSVMA	CPSCVCRDAGFYCNDRFLTSIPTGIPEDA	60
Db	1	MISPAWSLFLIGTKIGLFQVAPLSVMA	CPSCVCRDAGFYCNDRFLTSIPTGIPEDA	60
QY	61	TTLYLQNNQINNAGIPSDLNKLLKRYIYHNSLDEFPNLPKYKELHLOENNRIT	120	
Db	61	TTLYLQNNQINNAGIPSDLNKLLKRYIYHNSLDEFPNLPKYKELHLOENNRIT	120	
QY	121	YDLSKIPYELHLDONSYSVSIIEGAFRDSNLYRLFLSRNHLSTIPWGPRTIEEL	180	
Db	121	YDLSKIPYELHLDONSYSVSIIEGAFRDSNLYRLFLSRNHLSTIPWGPRTIEEL	180	
QY	181	RDDNRISTISPSLQIGTSIKRELVDGNLNNHGLDKVFFNVLNLTSLVRSNLSRAA	240	
Db	181	RDDNRISTISPSLQIGTSIKRELVDGNLNNHGLDKVFFNVLNLTSLVRSNLSRAA	240	
QY	241	PVNLPGTNLRKLYIQDNHINRVPNPAFSTYRLQYKELDMNSNNLSNLPQGFDDLNITQL	300	
Db	241	PVNLPGTNLRKLYIQDNHINRVPNPAFSTYRLQYKELDMNSNNLSNLPQGFDDLNITQL	300	
QY	301	ILRNPPWYCGCKMKWVRDWSLQSLPVKVNVRGLMCAQPEKVRGMAIKDLNAELFDCKDSGI	360	
Db	301	ILRNPPWYCGCKMKWVRDWSLQSLPVKVNVRGLMCAQPEKVRGMAIKDLNAELFDCKDSGI	360	

```
QY 361 VSTIQTITAPNTVYPAQOWPAPVTKQDPDKNPKLTKDOQTGSPGKTKITITVKSUTS 420
DB 361 VSIQITITAPNTVYPAQOWPAPVTKQDPDKNPKLTKDOQTGSPGKTKITITVKSUTP 420
QY 421 DTIHSKALPMTALRLSWLKGHPAFSGITETITVGERSEYLVATLPSDPYKVMV 480
DB 421 DTIHSKALPMTALRLSWLKGHPAFSGITETITVGERSEYLVATLPSDPYKVMV 480
QY 481 PMETSNIYLFDETPVCIETETAPLRYMNTPTTLNREQEKPKPKNPPLAALIGGAVLY 540
DB 481 PMETSNIYLFDETPVCIETETAPLRYMNTPTTLNREQEKPKPKNPPLAALIGGAVLY 540
QY 541 TIALALVWYVHRNGSLFSRNCAYSGRRKDDYAEAGTKKDSILFIRETSFQMLPIS 600
DB 541 SIALALVWYVHRNGSLFSRNCAYSGRRKDDYAEAGTKKDSILFIRETSFQMLPIS 600
QY 601 NEPISEEFVITIPPNGMNLKNNHSESSSNRSRSDSGIPDSDSHS 649
DB 601 NEPISEEFVITIPPNGMNLKNNHSESSSNRSRSDSGIPDSDSHS 649

RESULT 2
US-09-188-930-124
; Sequence 124, Application US/09188930A
; Patent No. 6150502
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James Greg
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; FILE REFERENCE: 11000.1011c1
; CURRENT APPLICATION NUMBER: US/09/188,930A
; CURRENT FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 348
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 124
; LENGTH: 110
; TYPE: PRT
; ORGANISM: mouse
US-09-188-930-124

Query Match 16.1%; Score 546; DB 4; Length 110;
Best Local Similarity 92.7%; Pred. No. 8.4e-42;
Matches 102; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 MISAAWSIFLIGKIGLFTQVAPLSVMAKSCPSVCRDAGFYCNDRLFTSIPTGIPEDA 60
DB 1 MISAPSEFLIGKIGLFTQVAPLSVMAKSCPSVCRDAGFYCNDRLFTSIPTGIPEDA 60
QY 61 TTYLQNNQINAGIPSDKLNLLKVERIYLYHNSLDEPFTNPKYVKELH 110
DB 61 TTYLQNNQINAGIPSDKLNLLKVERIYLYHNSLDEPFTNPKYVKELH 110

RESULT 3
US-09-131-648-2
; Sequence 2, Application US/09131648
; Patent No. 6168920
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Yue, Henry
; APPLICANT: Corley, Neil C.
; APPLICANT: Guegler, Karl J.
; APPLICANT: Patterson, Chandra
; TITLE OF INVENTION: EXTRACELLULAR ADHESIVE PROTEINS
; FILE REFERENCE: PF-0576 US
; CURRENT APPLICATION NUMBER: US/09/131,648
; CURRENT FILING DATE: 1998-08-10
; NUMBER OF SEQ ID NOS: 5
```

```
; SOFTWARE: PERL Program
; SEQ ID NO 2
; LENGTH: 708
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE: -
; OTHER INFORMATION: 2687731
US-09-131-648-2

Query Match 11.7%; Score 397; DB 4; Length 708;
Best Local Similarity 23.2%; Pred. No. 4.7e-27;
Matches 173; Conservative 114; Mismatches 240; Indels 218; Gaps 35;

QY 10 LIGKIGLFTQVAPLSVMAKSCPSVCRD-----DAGFYCNDRLFTSIPTGI 56
DB 11 LIGLATTIVQADKKV--DCPRLCTCEIRPWFPRSYMEASTVDCNDLGLLTPAKL 67
QY 57 PEDATLYLQNNQI-----NNAGIPSDKLNLLKVERI-----YLYHNS 94
DB 68 PANTQILLQTNNAKIEYSTDFPNLTGLDLSQNNLSVTNINVKMPOLLSVYLEENK 127
QY 95 LDEEP-----TNLPK-YVK-----ELHQENNIETIYDLSKI 127
DB 128 LTELPEKCLSELNLOELYINHLNSTISPGAFIGHLLRLHLNSRLQMINSKWFDAL 187
QY 128 PYLEELHLDNDSVSAVSTEGAFRDSNYLRLFLSLRNHLSTIP---WGLPRTIEELRD 183
DB 188 PNLLEILMIGENFI--IRKDMNFPLINLSLVIAGINLTETPDNALVGL-ENLESISY 244
QY 184 DNRISTISSPSLOGLTSILKRLVLGDNLLNHHGLGDKVPFNVLNLTSLVNSL-----237
DB 245 DNRLLIKYPHVALQKVNNLFLDLNKNPINIRRGD--FSNHLKELGI--NNMPELISI 300
QY 238 -TAAPVNLPGTNRKLYLQDN-HINRYPNAPSYLRLQYLDMSNNLSNLPDGLD 295
DB 301 DSLAVNLP--DLRKIEATNPRLSYTHPNAPFRPLKLESMLNSALSLYHGTIESLP 358
QY 296 NITQLILRNPNWYCGCKMKWYRDWLQSLPVKNVR-----GLMCOAPEKVRGMAIK--346
DB 359 NLKEISHSNIRCDVIRKWN-----MKNTRIMEPDSLCFVDPPEFGQNVQVHF 412
QY 347 -----DLNAEL-----PDKXDSIVSTIQ-----ITTA-----IPN 372
DB 413 RDMMEICLPLIAPSFPSNINVEAGSVVSPHCR-----ATAEPOPEIYWTTPSQKLLPN 467
QY 373 T-----VYPAQOWPAPVTKQ---PDINKELTKD---QQTGSPSR 408
DB 468 TLTDKTVHSEGTLDINGVTPREGGLYTCIATNLVGADLKSVMKVDGSPFDNNGS---524
QY 409 KTIITVKSVTSDTIHSKALPMTALRLSW---LKLGHSPAFSGITETITVGERSEYL 465
DB 525 --LMIKIRIQDQANSVLNSWKASSKILKSSVKWTAFAVKTENSHAAQSAR---IPSDVKVN 579
QY 466 VTALPEPSYKVMVPMETSNIYLFDETPVCIETETAPLRYMNTPTTLNREQEKPKPKNP 525
DB 580 LTHNPSTFKICI-----DIPTIYQKNEK--KCVNVTIKGLHPDQKEYEKN 625
QY 526 NLPLAALIGGAVLYTIALALVWYVHRNGSLFS--RNC-----AYSKGRERKDDYA--576
DB 626 TTYLMACLGGLLIGIVICLI-----SCLSPENMCDGSHSVYRNLYQKPTFAIGE 675
QY 577 -----EAGTKKDSILFIRET 592
DB 676 LYPPLINLWEAGKEKSTS-LKVKAT 699

RESULT 4
US-09-182-024A-2
; Sequence 2, Application US/09182024A
; Patent No. 6342370
; GENERAL INFORMATION:
; APPLICANT: Connolly, Timothy
; APPLICANT: Rajput, Bhanu
```

```

; Patent No. 6225085
; GENERAL INFORMATION:
; APPLICANT: Holtzman, Douglas A.
; TITLE OF INVENTION: NOVEL LRSG PROTEIN AND NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: MEI-019
; CURRENT APPLICATION NUMBER: US/09/063.950C
; CURRENT FILING DATE: 1998-04-21
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 673
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-063-950-2

```

Query Match	10.4%; Score 354.5; DB 4; Length 673;
Best Local Similarity	24.5%; Pred. No. 3e-23;
Matches 156;	Conservative 101; Mismatches 261; Indels 119; Gaps 261
QY	17 LFQVAPLSMAKSCPVGRC-DAGFIYCNDRFITSIPGIPEDATTYLQNNQI-----70
Db	13 LLLALGP---GVGCPSCOCQOPQVETCAQGTQVPRDVPDQTVGLYFENGITMLDA 69
QY	71 -NAGIPSDLKMLLKVERIYLHNSLDEFPPNPK-- -YVKELHLQNNRTITWYDSLSK 126
Db	70 GSYAGLGP-----LQILDSONQIASLSPGVQPLANSLDITANRLHETITNETFRG 122
QY	127 IPILELHLDKNSYSAVSEBAGFRDSNYLRLFLLSRNLHLSPI-PWGLPRTIEELRLDDN 185
Db	123 LRRLERYILGKNRIR--HIQGAFTIDRLLELELKQDNERALPRLRLPKL-----LILLDS 177
QY	186 RISTIS-SPSLOGITSLKRLVLDDGNLNNHGLGKVFNFVNLVLTSLVNSLTAAPVNL 244
Db	178 HNSLLALEPGILDANVEALRLAG--LGIQLODEGLFSRLRLNHLDDVSONQLERVPVI 235
QY	245 PG-TNLRLKLYLQDN-HLNRVPPNAPSRLQLRYLDDMSNNLSNLPQGLFDDLDNTQLIL 302
Db	236 RGLGRUTRLIAGNTRTAQURPBDLGAALQELDVSNLSUQLAPGLDLSGLFPRRLKLA 295
QY	303 RNNPWYCGCKMKVRDNLQSLPKV-NVRGLMCAQPKYRGMKIDNAELFCKDCKSGIV 361
Db	296 ARNPFNCVCLSPGWVWRSHVTLASPETTRCHFPFKNAGRLILLELDYADFQCPATTTT 355
QY	362 STQILI-----TALPNTVYPAQGNWAPYTKQPDIKNPKLT-----397
Db	356 ATWETTPVVRPEPTALSSLAFT---WLSPTAPATAPSPPTAPTPVGPVPOQDCPPS 412
QY	398-----KQDQTTGSPS-----RKTIITIVKSVT 419
Db	413 TCLNGGTCHLGRHRLACLCEGTGLYCESQMCGTSPSPPTVTPRPSRLSLGLTEPVS 472
QY	420 SDTHISWKLALPMTALRLSWLKLGHSPAFOSITETIVT---GERSEYLTALPDSPY 475
Db	473 PTLURLVGLQRYLQSSVOLRSLELYRNLGSP-DKRLVTLRLPDLASLAYTTLQRENTATY 531
QY	476 KVCMPWETSNNLYLDFDETPYCITETAPLRMYN--PTTTLNREQEKPKYKNPLPLAII 533
Db	532 SVCVPLGLGRVPEGER--ACGEAHTPPAVNSHAPVT-----QARE-----GNLPL--LI 578
QY	534 GGAAVLVTIALIAYL-CWYVHRNGSLFGRNCAYSKGR 569
Db	579 APALAAVTLAAALAAVGAAYCVRGRMAAA-AAODKGO 614

```

: GENERAL INFORMATION:
: APPLICANT: Connolly, Timothy
: APPLICANT: Rajput, Bhanu
: TITLE OF INVENTION: Human Slit Polypeptide and Polynucleotides Encoding

```

000 - - - FLEEFASSTFANUKRLDSSNNSIALADHADSUNQJLIL 373

QY 134 HLDNSVSAVSIEEAGAFRDSNYLRLLFLSRNLSTIPWGLPRTIE--LRDDNRISTI 190
| | : | : ||||| : | : | : | : | :
| | : | : ||||| : | : | : | : | :

191	SSPSLOGLTSLK----	202
QY	::: : :	
Db	434 ANGTFFDAMKSMKTVHLAKNPFFCOCNLKWLADYLHKNP1IETSGARCESPKRMHRR1ESL 493	
QY	203 -----BLVLDDGN----	211
Db	494 REEKFKCSWGLRDKMKSGBCRMDSCPAMCHCEGTVDCGTGRRLKXEIPRDIPLHTTLLI 553	
QY	212 NNHGLG----DKYFFENLVNLTSLSVRNSLTAAPVNL--PGNLRLKLYLQDNH1NRVPPN 265	
Db	554 NDNELGRISSDGLFGRPLHV4LEKRNQLTGTETPNAFEGASHIQLOIGENK1KE1SNK 613	
QY	266 AFSYLRQLYRLDMSNNLSNLPGQTFDDLD1N1TQ1LI1RNPWYCCGCKWKWYDWO1SLPV 325	
Db	614 MFLG1HQ1K1NL1NDQ1SCVMPGSEFHNLS1SLN1ASPNFNCN1HLAWFAECVRK-K 671	
QY	326 KYNVRGLMCQAEKVRGMAIKDLNLAELFDC 355	
Db	672 SLNGAARCGAPSKVRDQ1KDLPHSEFK 701	

QY 22 APLSVMAK-SCPSVCRCDAGFYICNDRFLTSTPTGIPEDATTLYLQNQINNAGIPSDLK 80
||| : ||| ||| | : | : ||| : | : | : | : |

Db 285 AMECCGAENSCPHRCADGIVDCREKSLTSVPVLPDDTTDVRLSQN----- 332
 QY 81 NLLKVERIYLYHNSIDEPNLP-----KYVKELHLOENNIKTITVDSLSKIPYLEL 133
 Db 333 -----FTELPKFSFSSFRRLRIDLSNNISRIAHDAISGLKQLTWL 375
 QY 134 HLDNSVSAVIEGAFRDSNYLRLLFLSRNHLSTIPWGLPRTIEE---LRLDNRISTI 190
 Db 376 VLYGNKIK--DLPSGVFKGLSLRLLLNANEISCIKRDAPDLHSLSLSDYDNNIQSL 433
 QY 191 SSPSLOGLSLK----- 202
 Db 434 ANGTEDAMKSMKTVHLAKNPFICDNLRLWADYLHKNPIETSGARCESPKRMHRRIESL 493
 QY 203 -----RLVLDGN-----LL 211
 Db 494 REEKFKCSWGEIRMKLSGECRMDSDCPAMCHCEGTTVCTGRRLKEIPRDIPLHTTELL 553
 QY 212 NNHGLG---DKVFFNLVNIETSLVRSNLTAAPVNL--PGTNLRKLYLQDNHNRVPPN 265
 Db 554 NDNELGRISSDGLFGRPLHLVLELKRNOITGIEPNAFEGASHIQELQGENKIKEISNK 613
 QY 266 AFSYLRQLYRLDMNNLSNLPQIFDDLDNITQLILRNPNWYCGCKMKWVRDWSLPV 325
 Db 614 MFLGLHQLKTLNLYDNOISCMFSGFSEHLNLSLTLASNFNCNHLAWFAECVRK--K 671
 QY 326 KVNVRGLMCOAPEKVRGMAIKDLNAELFDC 355
 Db 672 SLNGGAARCGAPSKVRDVQIKDLPHEFKC 701

RESULT 9
 US-09-540-153-7
 ; Sequence 7, Application US/09540153
 ; Patent No. 627095
 ; GENERAL INFORMATION:
 ; APPLICANT: Goodman, Corey
 ; APPLICANT: Kid, Thomas
 ; APPLICANT: Brose, Katja
 ; APPLICANT: Tessier-Lavigne, Marc
 ; TITLE OF INVENTION: Modulating Robo: Ligand Interactions
 ; FILE REFERENCE: B98-031-3
 ; CURRENT APPLICATION NUMBER: US/09/540,153
 ; CURRENT FILING DATE: 2000-03-31
 ; PRIOR APPLICATION NUMBER: 09/191,647
 ; PRIOR FILING DATE: 1998-11-13
 ; PRIOR APPLICATION NUMBER: 60/081,057
 ; PRIOR FILING DATE: 1998-04-07
 ; NUMBER OF SEQ ID NOS: 14
 ; SOFTWARE: Patent In Ver. 2.0
 ; SEQ ID NO 7
 ; TYPE: PRT
 ; LENGTH: 1480
 ; ORGANISM: Drosophila melanogaster
 US-09-540-153-7

Query Match 10.1%; Score 344.5; DB 4; Length 1480;
 Best Local Similarity 24.0%; Pred. No. 8.4e-22;
 Matches 108; Conservative 60; Mismatches 133; Indels 149; Gaps 10;
 QY 22 APLSVNAK--SCPSVCRDAGIYCNDRFTSTIPTGIPEDATLYLQNNQINNAGIPSLK 80
 Db 285 AMECCGAENSCPHRCADGIVDCREKSLTSVPVLPDDTTDVRLSQN----- 332
 QY 81 NLLKVERIYLYHNSIDEPNLP-----KYVKELHLOENNIKTITVDSLSKIPYLEL 133
 Db 333 -----FTELPKFSFSSFRRLRIDLSNNISRIAHDAISGLKQLTWL 375
 QY 134 HLDNSVSAVIEGAFRDSNYLRLLFLSRNHLSTIPWGLPRTIEE---LRLDNRISTI 190
 Db 376 VLYGNKIK--DLPSGVFKGLSLRLLLNANEISCIKRDAPDLHSLSLSDYDNNIQSL 433
 QY 191 SSPSLOGLSLK----- 202

Db 434 ANGTEDAMKSMKTVHLAKNPFICDNLRLWADYLHKNPIETSGARCESPKRMHRRIESL 493
 QY 203 -----RLVLDGN-----LL 211
 Db 494 REEKFKCSWGEIRMKLSGECRMDSDCPAMCHCEGTTVCTGRRLKEIPRDIPLHTTELL 553
 QY 212 NNHGLG---DKVFFNLVNIETSLVRSNLTAAPVNL--PGTNLRKLYLQDNHNRVPPN 265
 Db 554 NDNELGRISSDGLFGRPLHLVLELKRNOITGIEPNAFEGASHIQELQGENKIKEISNK 613
 QY 266 AFSYLRQLYRLDMNNLSNLPQIFDDLDNITQLILRNPNWYCGCKMKWVRDWSLPV 325
 Db 614 MFLGLHQLKTLNLYDNOISCMFSGFSEHLNLSLTLASNFNCNHLAWFAECVRK--K 671
 QY 326 KVNVRGLMCOAPEKVRGMAIKDLNAELFDC 355
 Db 672 SLNGGAARCGAPSKVRDVQIKDLPHEFKC 701

RESULT 10
 PCT-US91-09055-2
 ; Sequence 2, Application PC/US9109055
 ; GENERAL INFORMATION:
 ; APPLICANT: Rothberg, Jonathan Marc and Artavanis-Tsakonas, Spyridon
 ; TITLE OF INVENTION: Purified SLIT protein and Sequence Elements Thereof
 ; NUMBER OF SEQUENCES: 9
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Yale University
 ; ADDRESSEE: Office of Cooperative Research
 ; STREET: 246 Church Street
 ; CITY: New Haven
 ; STATE: Connecticut
 ; COUNTRY: USA
 ; ZIP: 06510
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette, 3.50 inch, 800 kb storage
 ; COMPUTER: Apple Macintosh
 ; OPERATING SYSTEM: Macintosh 6.0.5
 ; SOFTWARE: Microsoft Word 4.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/US91/09055
 ; FILING DATE: 19911127
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/624,135
 ; FILING DATE: 7-DEC-1990
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Barth, Richard J.
 ; REGISTRATION NUMBER: 28,180
 ; REFERENCE/DOCKET NUMBER: 900964/RB
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (212) 972-1400
 ; TELEFAX: (212) 370-1622
 ; TELEX: 236268
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1480 amino acids
 ; TYPE: AMINO ACIDS
 ; TOPOLOGY: Linear
 ; MOLECULE TYPE: protein
 ; FEATURE:
 ; NAME/KEY: signal sequence
 ; LOCATION: 1 to 36
 ; IDENTIFICATION METHOD: similarity to other signal
 ; OTHER INFORMATION: Directs Export
 ; NAME/KEY: Four Flank-LRR-Flank domains
 ; LOCATION: 37 to 910
 ; IDENTIFICATION METHOD: Array of Flank-LRR-Flank
 ; OTHER INFORMATION: mediates adhesive events
 ; NAME/KEY: tandem EGF-like repeats
 ; LOCATION: 911 to 1150

IDENTIFICATION METHOD: similarity to tandem EGF-like
OTHER INFORMATION: protein-protein interactions
NAME/KEY: 7th EGF-like repeat
LOCATION: 1353 to 1393
IDENTIFICATION METHOD: similarity to epidermal growth
OTHER INFORMATION: involvement in receptor-ligand
NAME/KEY: Alternative splice segment
LOCATION: 1394 to 1404
IDENTIFICATION METHOD: experimentally regulated
OTHER INFORMATION: developmentally regulated
NAME/KEY: COOH-terminal region
LOCATION: 1405 to 1480
IDENTIFICATION METHOD: experimental
PCT-US91-09055-2

Query Match 10.1%; Score 344.5; DB 5; Length 1480;
Best Local Similarity 24.0%; Pred. No. 8.4e-22;
Matches 108; Conservative 60; Mismatches 133; Indels 149; Gaps 10;

QY 22 APLSVMAK-SCPSVCRDAGFYCNDRELTSTIPGIPEDATTLYLQNNQINNAGIPSDLK 80
Db 285 APMECGAENSCHPCRCADGIVDCREKSLTSVPVTLPPDDTVDVLEQN----- 332
QY 81 NLLKVERIYHNSLDEFTNLP-----KYVELHLQNNIRITVDSLSKIPYLEEL 133
Db 333 -----FTELPKFSFRRRLRDLNNSNISRAHDALSGLKQLTTL 375
QY 134 HLDNSVSAYSIEGAFRDSNTLRLFLSRNHLSTIPWGLPTIEB---LRLDNRISTT 190
Db 376 VLYGNKIK--DLPSGVFKGLSGRLLLNANEISCIKDAFRDLHLSLSLYDNNQSL 433
QY 191 SPSPQGLTSLK----- 202
Db 434 ANGTFDAMKSMKTVHLAKNPFICDCNLRWLADYLRKNPIETSGARCESPKRMHRRTESL 493
QY 203 -----RLVLDGN----- 211
Db 494 REEFKFSWELRKLKSGCRMDSDCPAMCHCBGTVDCTGRKLEIPRDLPHTELL 553
QY 212 NNHGIG---DKVFFNLVNLTELSLVRSNLSLTAAPVNL--PGTNLRKLYLDNNHNVPPN 265
Db 554 NDELGRISDGLFGRPLHLVLEKRNQGLGIEPNAFEGASHIQELQGENKIKELSNK 613
QY 266 AFSYRLQYLRIDMSNNLSNLQGGIFDDLDNITOLIIRNNPWCCKMKWVDMQSLPV 325
Db 614 MFLGLHOLKTLNLYDQISCVMPGSEFHLNLSLNLASPNFCNCHLAWFAECVRK--K 671
QY 326 KYNVRGLMCOAPEKVRGMAIKDINAELEFDC 355
Db 672 SLNGGAARCAPSKYRDVQIKDLPHSEPKC 701

RESULT 11
US-09-191-647-2
Sequence 2, Application US/09191647
Patent No. 6046015
GENERAL INFORMATION:
APPLICANT: Goodman, Corey
APPLICANT: Kid, Thomas
APPLICANT: Brose, Katja
APPLICANT: Tessier-Lavigne, Marc
TITLE OF INVENTION: Modulating Robo: Ligand Interactions
FILE REFERENCE: B98-031-3
CURRENT APPLICATION NUMBER: US/09/191,647
CURRENT FILING DATE: 1998-11-13
EARLIER APPLICATION NUMBER: 60/065,544
EARLIER FILING DATE: 1997-11-14
EARLIER APPLICATION NUMBER: 60/081,057
EARLIER FILING DATE: 1998-04-07
NUMBER OF SEQ ID NOS: 14
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 1525

TYPE: PRT
ORGANISM: human
US-09-191-647-2

Query Match 10.1%; Score 343.5; DB 3; Length 1525;
Best Local Similarity 23.9%; Pred. No. 1.1e-21;
Matches 119; Conservative 55; Mismatches 165; Indels 159; Gaps 13;

QY 1 MISAAWSIFELGKIGLFL---QVAPLSVMAKSCPSVCRDAGFYCNDRELTSTIPGI 56
Db 1 MRGVGWQ--MISLSGLVLAAILNKVAP-----QACPAQSCSGSTVDCGHLALRSVPRNI 53
QY 57 PEDATLYLQNNQIN-----NAGIPSDLKMLKVERIYL 90
Db 54 PRNTERLDLNGNITRITKTDFAGLRHLVQLMENKISTIERGAFQDKEL---ERLRL 110
QY 91 YHNSLDEFPNL---PKYKEHLQNNIRITVDSLSKIPYLEELHDDNSVAVSIEE 147
Db 111 NRNHLQFLPELLEFLGAKLYRLDLSENQIOAIPKAPGAVDIKNLQLDYQISC--IED 168
QY 148 GAFR-----DSNYLRLELFLSRNH----- 165
Db 169 GAFRALRDLEVLFINNNNITRLSVASFNMHPKLTFFLHNSNLYCDCHLAWLSDWLKRRP 228
QY 166 ----- 165
Db 229 RYGLYTCMGPSHLRGHNAEYOKREFVCSDEERGHQSFMAPSCSVLHCPAACSNNIV 288
QY 166 -----LSTIPWGLPTIEBELDDNRISTISSPSQGLTSLKRLVLDGNLNNHGLDK 219
Db 289 DCRGKGLTEIPTNLPTETIETLEQNTIKVPPGAFSPYKFLERIDLNNQISE--LAPD 346
QY 220 VFENLVNLTELSLVRSNLSLTAAPVNL--PGTNLRKLYLDNNHNVPPNPAFSLYRLQYRLD 277
Db 347 AFOGLRSLSLVLYGKNIETELPKSLFEGLSQLLLNANKINCIEVDAFQDLHNLNLS 406
QY 278 MSNNLSNLQGGIFDDLDNITOLIIRNNPWCCKMKWVDMQSLPVKVNVRGLMCOAP 337
Db 407 LYDNKLOTIARTGTSPSLRAIOTMHLAQNPFCDCCHLAWLADYLRHTNPTETS--GARCTSP 464
QY 338 EKVRGMAIKDINAELEFDC 355
Db 465 RELANKRIGQIKSKKPRC 482

RESULT 12
US-09-540-245A-2
Sequence 2, Application US/09540245A
Patent No. 6270984
GENERAL INFORMATION:
APPLICANT: Goodman, Corey
APPLICANT: Kid, Thomas
APPLICANT: Brose, Katja
APPLICANT: Tessier-Lavigne, Marc
TITLE OF INVENTION: Modulating Robo: Ligand Interactions
FILE REFERENCE: B98-031-3
CURRENT APPLICATION NUMBER: US/09/540,245A
CURRENT FILING DATE: 2000-03-31
PRIOR APPLICATION NUMBER: 60/065,544
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: 60/081,057
PRIOR FILING DATE: 1998-04-07
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 1525
TYPE: PRT
ORGANISM: human
US-09-540-245A-2

Query Match 10.1%; Score 343.5; DB 4; Length 1525;
Best Local Similarity 23.9%; Pred. No. 1.1e-21;
Matches 119; Conservative 55; Mismatches 165; Indels 159; Gaps 13;

[illegible]

```

RESULT 13
US-09-540-153-2
; Sequence 2, Application US/09540153
; Patent No. 6270995
; GENERAL INFORMATION:
; APPLICANT: Goodman, Corey
; APPLICANT: Kid, Thomas
; APPLICANT: Brose, Katja
; APPLICANT: Tessier-Lavigne, Marc
; TITLE OF INVENTION: Modulating Robo: Ligand Interactions
; FILE REFERENCE: B98-031-3
; CURRENT APPLICATION NUMBER: US/09/540,153
; CURRENT FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/191,647
; PRIOR FILING DATE: 1998-11-13
; PRIOR APPLICATION NUMBER: 60/081,057
; PRIOR FILING DATE: 1998-04-07
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1525
; TYPE: PRT
; ORGANISM: human
US-09-540-153-2

```

```

Db      54  PNTYERLDLNGNNIRIYKTFAGLRLHRLVQLQIMENKISTIERGAFQDLKEL---ERLRL 111
QY      91  YHNSLDEPFTNL---PKYVKELHLQENNIRITITYDSLKSPIYLEELHLDSDSVSAVSTIEE 147
Db      111  NNNHLQLFPELLFLGTAKLYRLDLSNQIQATPKAFTGAVDIKNLQLDYNOISC--IED 168
QY      148  GAFR-----DSVNLRLFLSRNH-----165
Db      169  GAFRALRDLEVLFLNANNITSLVSASFNNHMPKLRTRFLHNSNLYCDBLAWLSDWLKRKP 228
QY      166  -----165
Db      229  RVGLYTQMGSPSLRGHNVAVQKREVVSDEEGHQSFMAPSCSVHCPAATCSNNIV 288
QY      166  -----LSTIPWGLPRTIEELRLDDNRKISTISPSLQGLTSLKRLVLDGNLNNHGLGDK 219
Db      289  DCRGKGLTEIPTNLPTETIEIRLQNTIKVPPGAFSPYKKLRRLDSNNQISE--LAPD 346
QY      220  VFFNLVNLTELSLVNRSLTAAPVNL-PGT-NURKLYLDNHNIRVPNAPSVLRLQLRYLD 277
Db      347  APOGLRSINSVLVYGNKITEPKSLFGLFSLQLLLLNANKINCLRVDAFQDLHNLMLS 406
QY      278  MNNNLSNLPGQIFDDLDLNIQTLLLRNNWYCGCKMKWVRDLQSLPKYVNVVRGLMQAP 337
Db      407  LYDNKLQTIANGFSPSLRAITMELAQNFICDCHKWLDYILHTNPIETS--GARTCP 464
QY      338  EKVGRMAIKDLNAELFDC 355
Db      465  RELANKRIGQIKSKKFC 482

RESULT 14
US-08-442-063A-45
; Sequence 45, Application US/08442063A
; Patent No. 5705609
; GENERAL INFORMATION:
; APPLICANT: RUOSLAHTI, ERKKI I.
; APPLICANT: PIERSCHEBACHER, MICHAEL D.
; APPLICANT: CARDENAS, JOSE
; APPLICANT: CRAIG, WILLIAM
; APPLICANT: MULLEN, DANIEL G.
; TITLE OF INVENTION: DECORIN FRAGMENTS AND METHODS OF
; TITLE OF INVENTION: INHIBITING CELL REGULATORY FACTORS
; NUMBER OF SEQUENCES: 62
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CAMPBELL AND FLORES
; STREET: 4370 LA JOLLA VILLAGE DRIVE, SUITE 700
; CITY: SAN DIEGO
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/442,063A
; FILING DATE: 16-MAY-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/865,652
; FILING DATE: 03-APR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: CAMPBELL, CATHRYN A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LA 1454
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-535-9001
; TELEFAX: 619-535-8949
; INFORMATION FOR SEQ ID NO: 45:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 282 amino acids

```

QY 57 PEDATTIYLQNNQIN-----NAGIPSDLKNLKVERIYL 90
| : | | | | | | | | | | | | | | | | | |

TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-442-063A-45

Query Match 10.1%; Score 342; DB 1; Length 282;
Best Local Similarity 31.9%; Pred. No. 9.7e-23;
Matches 90; Conservative 49; Mismatches 85; Indels 58; Gaps 9;

QY 31 CPSCVCRDAGFYCNDRLTSLPTGIPEDATTTLYLQNNQINNA--GIPSLKN-----81
DB 26 CPFCQCHLVVQVCSDLGDKVPKDLPPDFTLLDQNNKITEIKDGFKNLKNHALILV 85
QY 82 -----LLKVERIYLYHNSLDEFPNLPKYVKELHLOENNI---RTIYDSL 124
DB 86 NNKISKVSPCAFTPLVKLERLYLSKNQKELPEKMEKTIQELRAHENEITKVKVTFNGI 145
QY 125 SKIPLYELHLDNSVSAVTEEGAFRDSNYLRLFLSRNHLSTIPWGLPRTTEELRLDD 184
DB 146 NQMTIVIE--LGTNPLKSGIENGAFQGMKKLSYIRIADTNTITSPQGLPESLTELHLDG 202
QY 185 NRISTISSPSLQGLTSLKRLVLDGNLNNHGLGDKVFFNVLNLTSLSVRNSLTAAPVNL 244
DB 203 NKISRVDAASLKGNNLAKLGLSPN-----SISAVDNGSLA---NT 240
QY 245 PGTNLRKLYLDNHNIRVPPN--AFSYLRQLYRLDMSNNLS 284
DB 241 P--HLRELHLDNNKLTFRVPGGLAEHKYIQVY---LHNNNIS 277

RESULT 15
US-08-442-063A-48
Sequence 48, Application US/08442063A
Patent No. 5705609
GENERAL INFORMATION:
APPLICANT: RUOSLAHTI, ERKKI I.
APPLICANT: PERSCHBACHER, MICHAEL D.
APPLICANT: CARDENAS, JOSE
APPLICANT: CRAIG, WILLIAM
APPLICANT: MULLEN, DANIEL G.
TITLE OF INVENTION: DECORIN FRAGMENTS AND METHODS OF
TITLE OF INVENTION: INHIBITING CELL REGULATORY FACTORS
NUMBER OF SEQUENCES: 62
CORRESPONDENCE ADDRESS:
ADDRESSEE: CAMPBELL AND FLORES
STREET: 4370 LA JOLLA VILLAGE DRIVE, SUITE 700
CITY: SAN DIEGO
STATE: CALIFORNIA
COUNTRY: UNITED STATES
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/442,063A
FILING DATE: 16-MAY-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/865,652
FILING DATE: 03-APR-1992
ATTORNEY/AGENT INFORMATION:
NAME: CAMPBELL, CATHRYN A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-LA 1454
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-535-9001
TELEFAX: 619-535-8949
INFORMATION FOR SEQ ID NO: 48:
SEQUENCE CHARACTERISTICS:
LENGTH: 307 amino acids
TYPE: amino acid

TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-442-063A-48

Query Match 10.1%; Score 342; DB 1; Length 307;
Best Local Similarity 31.9%; Pred. No. 1.1e-22;
Matches 90; Conservative 49; Mismatches 85; Indels 58; Gaps 9;

QY 31 CPSCVCRDAGFYCNDRLTSLPTGIPEDATTTLYLQNNQINNA--GIPSLKN-----81
DB 26 CPFCQCHLVVQVCSDLGDKVPKDLPPDFTLLDQNNKITEIKDGFKNLKNHALILV 85
QY 82 -----LLKVERIYLYHNSLDEFPNLPKYVKELHLOENNI---RTIYDSL 124
DB 86 NNKISKVSPCAFTPLVKLERLYLSKNQKELPEKMEKTIQELRAHENEITKVKVTFNGI 145
QY 125 SKIPLYELHLDNSVSAVTEEGAFRDSNYLRLFLSRNHLSTIPWGLPRTTEELRLDD 184
DB 146 NQMTIVIE--LGTNPLKSGIENGAFQGMKKLSYIRIADTNTITSPQGLPESLTELHLDG 202
QY 185 NRISTISSPSLQGLTSLKRLVLDGNLNNHGLGDKVFFNVLNLTSLSVRNSLTAAPVNL 244
DB 203 NKISRVDAASLKGNNLAKLGLSPN-----SISAVDNGSLA---NT 240
QY 245 PGTNLRKLYLDNHNIRVPPN--AFSYLRQLYRLDMSNNLS 284
DB 241 P--HLRELHLDNNKLTFRVPGGLAEHKYIQVY---LHNNNIS 277

Search completed: April 23, 2003, 11:07:14
Job time : 22 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 11:06:26 ; Search time 22 Seconds
(without alignments)
2363.829 Million cell updates/sec

Title: US-10-004-551-6
Perfect score: 3400
Sequence: 1 MISAAWSIFLTGKIGLFQ.....SSNSRYSRDSGIPDSHSHS 649

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 301932 seqs, 80129803 residues

Total number of hits satisfying chosen parameters: 301932

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : Published Applications AA.*
- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
 - 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
 - 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
 - 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
 - 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
 - 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
 - 7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
 - 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
 - 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
 - 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
 - 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
 - 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
 - 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
 - 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	3400	100.0	649	US-10-063-547-132	Sequence 132, App
2	3400	100.0	649	US-10-004-551-6	Sequence 6, Appl
3	3400	100.0	649	US-10-174-590-384	Sequence 384, App
4	3400	100.0	649	US-10-176-758-384	Sequence 384, App
5	3400	100.0	649	US-10-063-616-132	Sequence 132, App
6	3400	100.0	649	US-10-175-737-384	Sequence 384, App
7	3400	100.0	649	US-10-063-502-132	Sequence 132, App
8	3400	100.0	649	US-10-173-706-384	Sequence 384, App
9	3400	100.0	649	US-10-175-738-384	Sequence 384, App
10	3400	100.0	649	US-10-175-752-384	Sequence 384, App
11	3400	100.0	649	US-10-176-482-384	Sequence 384, App
12	3400	100.0	649	US-10-176-757-384	Sequence 384, App
13	3400	100.0	649	US-10-176-913-384	Sequence 384, App
14	3400	100.0	649	US-10-180-552-384	Sequence 384, App
15	3400	100.0	649	US-10-180-552-384	Sequence 384, App
16	3400	100.0	649	US-10-173-700-384	Sequence 384, App
17	3400	100.0	649	US-10-173-700-384	Sequence 384, App
18	3400	100.0	649	US-10-174-572-384	Sequence 384, App
19	3400	100.0	649	US-10-174-582-384	Sequence 384, App

20	3400	100.0	649	US-10-174-588-384	Sequence 384, App
21	3400	100.0	649	US-10-175-739-384	Sequence 384, App
22	3400	100.0	649	US-10-175-740-384	Sequence 384, App
23	3400	100.0	649	US-10-175-743-384	Sequence 384, App
24	3400	100.0	649	US-10-176-488-384	Sequence 384, App
25	3400	100.0	649	US-10-176-492-384	Sequence 384, App
26	3400	100.0	649	US-10-176-747-384	Sequence 384, App
27	3400	100.0	649	US-10-176-750-384	Sequence 384, App
28	3400	100.0	649	US-10-176-985-384	Sequence 384, App
29	3400	100.0	649	US-10-176-987-384	Sequence 384, App
30	3400	100.0	649	US-10-176-991-384	Sequence 384, App
31	3400	100.0	649	US-10-176-992-384	Sequence 384, App
32	3400	100.0	649	US-10-176-993-384	Sequence 384, App
33	3400	100.0	649	US-10-184-658-384	Sequence 384, App
34	3400	100.0	649	US-10-173-695-384	Sequence 384, App
35	3400	100.0	649	US-10-173-697-384	Sequence 384, App
36	3400	100.0	649	US-10-173-705-384	Sequence 384, App
37	3400	100.0	649	US-10-174-576-384	Sequence 384, App
38	3400	100.0	649	US-10-174-585-384	Sequence 384, App
39	3400	100.0	649	US-10-174-586-384	Sequence 384, App
40	3400	100.0	649	US-10-175-747-384	Sequence 384, App
41	3400	100.0	649	US-10-176-481-384	Sequence 384, App
42	3400	100.0	649	US-10-176-485-384	Sequence 384, App
43	3400	100.0	649	US-10-176-487-384	Sequence 384, App
44	3400	100.0	649	US-10-176-493-384	Sequence 384, App
45	3400	100.0	649	US-10-176-756-384	Sequence 384, App

ALIGNMENTS

RESULT 1

US-10-063-547-132
Sequence 132, Application US/10063547
Publication No. US20020182638A1

GENERAL INFORMATION:

- APPLICANT: Eaton, Dan L.
- APPLICANT: Filvaroff, Ellen
- APPLICANT: Gerritsen, Mary E.
- APPLICANT: Goddard, Audrey
- APPLICANT: Godowski, Paul J.
- APPLICANT: Grimaldi, Christopher J.
- APPLICANT: Gurney, Austin L.
- APPLICANT: Watanabe, Colin K.
- APPLICANT: Wood, William I.

TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

TITLE OF INVENTION: ACIDS ENCODING THE SAME

FILE REFERENCE: P3230R1C1

CURRENT APPLICATION NUMBER: US/10/063,547

Prior Filing Date: 2002-05-02

NUMBER OF SEQ ID NOS: 170

SEQ ID NO 132

LENGTH: 649

TYPE: PRT

ORGANISM: Homo Sapien

US-10-063-547-132

Query Match 100.0%; Score 3400; DB 9; Length 649;

Best Local Similarity 100.0%; Pred. No. 3.3e-261; Indels 0; Gaps 0;

Matches 649; Conservative 0; Mismatches 0;

1 MISAAWSIFLTGKIGLFQVAPLSVMAKSCFVRCDCAGFYCNDRFLTSIPTGIPEDA 60

|||||

1 MISAAWSIFLTGKIGLFQVAPLSVMAKSCFVRCDCAGFYCNDRFLTSIPTGIPEDA 60

|||||

61 TTYLQNNQINNAGIPSDLNKLVKVERIYIYNSLDEPTNLPKYVKELHLENNIRIT 120

|||||

61 TTYLQNNQINNAGIPSDLNKLVKVERIYIYNSLDEPTNLPKYVKELHLENNIRIT 120

|||||

121 YDSLKIPYVEELHLDNDSVNSIEGAFRDSNYLRLFLSRNHLSTIPWGLPTIEEL 180

|||||

121 YDSLKIPYVEELHLDNDSVNSIEGAFRDSNYLRLFLSRNHLSTIPWGLPTIEEL 180

|||||

Handwritten signature

QY	181	RIDNRRISTTSPTSOGTSLKRUVLDDGNLNHHGIDGKVFVNLVNLTELSVRNSTAA	240
DB	181	RUDNRRISTTSPSLQGLTSKRUVLDDGNLNHHGIDGKVFVNLVNLTELSVRNSTAA	240
QY	241	PVNLPGTNLRKLYLDONHINRVPNAPSYRLQYRLDMSNNLSNLPQGFDDLDNLTQL	300
DB	241	PVNLPGTNLRKLYLDONHINRVPNAPSYRLQYRLDMSNNLSNLPQGFDDLDNLTQL	300
QY	301	ILRNNPYCCKMKWYRWLQSLPVKVNVYRGLMCOAPEKVRGMAIKDLNAEFLPCKDSGI	360
DB	301	ILRNNPYCCKMKWYRWLQSLPVKVNVYRGLMCOAPEKVRGMAIKDLNAEFLPCKDSGI	360
QY	361	VSTQITAITPNVYPAQOWPAPVTKQPDINKPLTKDQQTGSPSRKTIITIVKSVTS	420
DB	361	VSTQITAITPNVYPAQOWPAPVTKQPDINKPLTKDQQTGSPSRKTIITIVKSVTS	420
QY	421	DTIHHSKWLALPMTALRLSNLKLGHSPAGSGTETIVTGERSEYILVTALEPDPSPKVCW	480
DB	421	DTIHHSKWLALPMTALRLSNLKLGHSPAGSGTETIVTGERSEYILVTALEPDPSPKVCW	480
QY	481	PMETSNIYLFDEPVCITETAPLRMYNPTTILNREQEKPEYKNPNLPLAAIIGGAVLV	540
DB	481	PMETSNIYLFDEPVCITETAPLRMYNPTTILNREQEKPEYKNPNLPLAAIIGGAVLV	540
QY	541	TIALLALVCWYHRNGSLFSBNCAYSKGRRRKDDYAEAGTKKONSILEIRETSFQMLPIS	600
DB	541	TIALLALVCWYHRNGSLFSBNCAYSKGRRRKDDYAEAGTKKONSILEIRETSFQMLPIS	600
QY	601	NEPTSKKEFVIHTIPPPGNMNYKNHSSSSNRYSRDSGIPDSHSHS	649
DB	601	NEPTSKKEFVIHTIPPPGNMNYKNHSSSSNRYSRDSGIPDSHSHS	649

```

RESULT 2
US-10-004-551-6
; Sequence 6, Application US/10004551
; Publication No. US20030004310A1
; GENERAL INFORMATION:
; APPLICANT: SHIMKETS, RICHARD A
; APPLICANT: FERNANDES, ELMA
; TITLE OF INVENTION: POLYPEPTIDES AND POLYPEPTIDES ENCODED THEREBY
; FILE REFERENCE: 15966-559
; CURRENT APPLICATION NUMBER: US/10/004,551
; CURRENT FILING DATE: 2001-12-05
; PRIOR APPLICATION NUMBER: 09/635,949
; PRIOR FILING DATE: 2000-08-10
; NUMBER OF SEQ ID NOS: 110
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 649
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-004-551-6

```

```
Query Match      100.0%; Score 3400; DB 9; Length 649;
Best Local Similarity 100.0%; Pred. No. 3.3e-361;
Matches 649; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

[illegible]

Db	181	RDDNRKTISSPSLQGLTSKLRLVLDGNLNNHGLGDKVFNVLWLTSLSVRNSLTAA	240
QY	241	PVNLPGTNRKLKYLQDNHINRPPNAPSYLRQLYRLDMSNNNSLNPQGFDDLDNLTQI	300
Db	241	PVNLPGTNRKLKYLQDNHINRPPNAPSYLRQLYRLDMSNNNSLNPQGFDDLDNLTQI	300
QY	301	ILRNPWTGCGKKKWRDMLQSLPVKYNVREGLMCOAPKRVGMAIKDLNAELFDCKDSGI	360
Db	301	ILRNPWTGCGKKKWRDMLQSLPVKYNVREGLMCOAPKRVGMAIKDLNAELFDCKDSGI	360
QY	361	VSTIQITTAIPNTVTPAQOWPAPVTKQPIDKNPKLTKDQQTGSPSRKTIITIVKSVTS	420
Db	361	VSTIQITTAIPNTVTPAQOWPAPVTKQPIDKNPKLTKDQQTGSPSRKTIITIVKSVTS	420
QY	421	DTIHISWKLALPMTALRLSWLKLGHSPAGSITETIVTGERSEVLYLTALPDSPTYKCMV	480
Db	421	DTIHISWKLALPMTALRLSWLKLGHSPAGSITETIVTGERSEVLYLTALPDSPTYKCMV	480
QY	481	PMETSNLYLDFEPPVCIETETAPLRMTNPTTLINRQSKPEYKPNPLPLAAIIGGVALV	540
Db	481	PMETSNLYLDFEPPVCIETETAPLRMTNPTTLINRQSKPEYKPNPLPLAAIIGGVALV	540
QY	541	TIALLALVCWVTHRNGSLFSRNCAYSGRRRKDDYAEAGTKKNSILEITRETSFQMLPIS	600
Db	541	TIALLALVCWVTHRNGSLFSRNCAYSGRRRKDDYAEAGTKKNSILEITRETSFQMLPIS	600
QY	601	NPISKBEFVLIHTFPPNGMNLKNNHSESSNRSRYDSGIPQSDSHS	649
Db	601	NPISKBEFVLIHTFPPNGMNLKNNHSESSNRSRYDSGIPQSDSHS	649

RESULT 3
US-10-174-590-384
; Sequence 384, Application US/10174590
; Publication No. US20030008352A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pau, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C42
; CURRENT APPLICATION NUMBER: US/10/174,590
; CURRENT FILING DATE: 2002-06-18
; Prior application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 384
; LENGTH: 649
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-174-590-384

	Query Match	100.0%;	Score 3400;	DB 9;	Length 649;
	Best Local Similarity	100.0%;	Pred. No. 3.3e-261;		
	Matches 649;	Conservative	0;	Mismatches 0;	Indels 0;
QY	1	MISAANSPIFGTKIGLFQVAPLSVMAKSPSCVCRDAGFIYCNDRFITSIPTGIPEDA	60		
Db	1	MISAANSPIFGTKIGLFQVAPLSVMAKSPSCVCRDAGFIYCNDRFITSIPTGIPEDA	60		
QY	61	TTLYLQNNQINNAGIPSDLNKLKVKRYIYLHNSLDEFPNLPKYVKEHLHQENNITIT	120		
Db	61	TTLYLQNNQINNAGIPSDLNKLKVKRYIYLHNSLDEFPNLPKYVKEHLHQENNITIT	120		
QY	121	YDSLSIKPIYELHLDONSAYSIEGAFRDSNYLLFLSRNHLSTITPWCGLPRTIEEL	180		

```

Db 121 YDSLSKIPIYEEHLDDNSVSAVSEEGAFRDSNYLRLLEFLSRNHLSTIPWGLPTIEEL 180
QY 181 RLDNRIISTSSPSLOGLTSLKRLVLDGNLNNHGLGDKVFFNLVNLTELSVRNSLTAA 240
Db 181 RLDNRIISTSSPSLOGLTSLKRLVLDGNLNNHGLGDKVFFNLVNLTELSVRNSLTAA 240
QY 241 PVNLPGTNLRLKYLQDNHINRVPPNFAFSLRQLYRLDMSNNLSNLPQGFDDLDNITQL 300
Db 241 PVNLPGTNLRLKYLQDNHINRVPPNFAFSLRQLYRLDMSNNLSNLPQGFDDLDNITQL 300
QY 301 ILRNNPWYCGCKMKWVRLQSLPVKVNVRGLMCAPEKVRGMAIKDLNAELFDCDGGI 360
Db 301 ILRNNPWYCGCKMKWVRLQSLPVKVNVRGLMCAPEKVRGMAIKDLNAELFDCDGGI 360
QY 361 VSTIQTITAIPTNTVYPAQGWPAFVTKOPDIKNPKLTKDOQTGSPSRKTITITVKSPTS 420
Db 361 VSTIQTITAIPTNTVYPAQGWPAFVTKOPDIKNPKLTKDOQTGSPSRKTITITVKSPTS 420
QY 421 DTHISWKLALPMTALRLSWLKLGHSPAFGSTTETITVGERSEYLVTALEPSPYKVCV 480
Db 421 DTHISWKLALPMTALRLSWLKLGHSPAFGSTTETITVGERSEYLVTALEPSPYKVCV 480
QY 481 PMETSNLYLDETTPVCTETETAPLRMYNPTTLNREQEKPEYKKNPLPLAAIIGGAVLV 540
Db 481 PMETSNLYLDETTPVCTETETAPLRMYNPTTLNREQEKPEYKKNPLPLAAIIGGAVLV 540
QY 541 TIALALVCWTVVHRNGLSFRNCAYSKGRRKDDYAEAGTKKNSILEIRETSFQMLPIS 600
Db 541 TIALALVCWTVVHRNGLSFRNCAYSKGRRKDDYAEAGTKKNSILEIRETSFQMLPIS 600
QY 601 NEPISKEEFVHTITFPNGMNLKNNHSESSNSRSDGIPDSHSHS 649
Db 601 NEPISKEEFVHTITFPNGMNLKNNHSESSNSRSDGIPDSHSHS 649

RESULT 4
US-10-176-758-384
; Sequence 384, Application US/10176758
; Publication No. US20030008353A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430RJC104
; CURRENT FILING DATE: 2002-06-21
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 384
; LENGTH: 649
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-176-758-384
Query Match 100.0%; Score 3400; DB 9; Length 649;
Best Local Similarity 100.0%; Pred. No. 3.3e-261;
Matches 649; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MISAWSIFLIGTKIGLFLQVAPLSVMAKSCPSVCRDAGFYCNDRFLTSTIPTGPEDA 60
Db 1 MISAWSIFLIGTKIGLFLQVAPLSVMAKSCPSVCRDAGFYCNDRFLTSTIPTGPEDA 60

```

```

QY 61 TTYLQNNQINNAGIPSDLKNLLKVERIYLYHNSLDEPTNLPKYVKELHLQENNRIT 120
Db 61 TTYLQNNQINNAGIPSDLKNLLKVERIYLYHNSLDEPTNLPKYVKELHLQENNRIT 120
QY 121 YDSLSKIPIYEEHLDDNSVSAVSEEGAFRDSNYLRLLEFLSRNHLSTIPWGLPTIEEL 180
Db 121 YDSLSKIPIYEEHLDDNSVSAVSEEGAFRDSNYLRLLEFLSRNHLSTIPWGLPTIEEL 180
QY 181 RLDNRIISTSSPSLOGLTSLKRLVLDGNLNNHGLGDKVFFNLVNLTELSVRNSLTAA 240
Db 181 RLDNRIISTSSPSLOGLTSLKRLVLDGNLNNHGLGDKVFFNLVNLTELSVRNSLTAA 240
QY 241 PVNLPGTNLRLKYLQDNHINRVPPNFAFSLRQLYRLDMSNNLSNLPQGFDDLDNITQL 300
Db 241 PVNLPGTNLRLKYLQDNHINRVPPNFAFSLRQLYRLDMSNNLSNLPQGFDDLDNITQL 300
QY 301 ILRNNPWYCGCKMKWVRLQSLPVKVNVRGLMCAPEKVRGMAIKDLNAELFDCDGGI 360
Db 301 ILRNNPWYCGCKMKWVRLQSLPVKVNVRGLMCAPEKVRGMAIKDLNAELFDCDGGI 360
QY 361 VSTIQTITAIPTNTVYPAQGWPAFVTKOPDIKNPKLTKDOQTGSPSRKTITITVKSPTS 420
Db 361 VSTIQTITAIPTNTVYPAQGWPAFVTKOPDIKNPKLTKDOQTGSPSRKTITITVKSPTS 420
QY 421 DTHISWKLALPMTALRLSWLKLGHSPAFGSTTETITVGERSEYLVTALEPSPYKVCV 480
Db 421 DTHISWKLALPMTALRLSWLKLGHSPAFGSTTETITVGERSEYLVTALEPSPYKVCV 480
QY 481 PMETSNLYLDETTPVCTETETAPLRMYNPTTLNREQEKPEYKKNPLPLAAIIGGAVLV 540
Db 481 PMETSNLYLDETTPVCTETETAPLRMYNPTTLNREQEKPEYKKNPLPLAAIIGGAVLV 540
QY 541 TIALALVCWTVVHRNGLSFRNCAYSKGRRKDDYAEAGTKKNSILEIRETSFQMLPIS 600
Db 541 TIALALVCWTVVHRNGLSFRNCAYSKGRRKDDYAEAGTKKNSILEIRETSFQMLPIS 600
QY 601 NEPISKEEFVHTITFPNGMNLKNNHSESSNSRSDGIPDSHSHS 649
Db 601 NEPISKEEFVHTITFPNGMNLKNNHSESSNSRSDGIPDSHSHS 649

RESULT 5
US-10-063-616-132
; Sequence 132, Application US/10063616
; Publication No. US20030013855A1
; GENERAL INFORMATION:
; APPLICANT: Baton, Dan L.
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3230R1C1
; CURRENT FILING DATE: 2002-05-03
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 132
; LENGTH: 649
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-063-616-132
Query Match 100.0%; Score 3400; DB 9; Length 649;
Best Local Similarity 100.0%; Pred. No. 3.3e-261;
Matches 649; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MISAWSIFLIGTKIGLFLQVAPLSVMAKSCPSVCRDAGFYCNDRFLTSTIPTGPEDA 60

```

```

Db 1 MISAWIFLIGKIGLFLQVAPLSVMAKSCPSVCRDAGFIYCNDRFLTSTPGIPEDA 60
QY 61 TTYLQNNQNNAGIPSDKLNLLKVERIYLYHNSLDEFFNLPKYKELHLOENNITIT 120
Db 61 TTYLQNNQNNAGIPSDKLNLLKVERIYLYHNSLDEFFNLPKYKELHLOENNITIT 120
QY 121 YDSLKIPYLEELHLDNDSVSAVSEEGAFRDSNYLLFLSRNHLSTIPWGLPRTTEEL 180
Db 121 YDSLKIPYLEELHLDNDSVSAVSEEGAFRDSNYLLFLSRNHLSTIPWGLPRTTEEL 180
QY 181 RLDNRISTSSSLOGLISLRVLDGNLNNHGLGDKVFFNVLNLTSLVRSNSTAA 240
Db 181 RLDNRISTSSSLOGLISLRVLDGNLNNHGLGDKVFFNVLNLTSLVRSNSTAA 240
QY 241 PVNLPFTNLRLKLYLDNHNRPVPPNAFSLRYLRDMSNNLSNLPQIFDDLDNITQL 300
Db 241 PVNLPFTNLRLKLYLDNHNRPVPPNAFSLRYLRDMSNNLSNLPQIFDDLDNITQL 300
QY 301 ILRNPWYCGCKMKWYRDLQSLPVKVNVRGLMQAPEKVRGMAIKDLNAELFCKDSGI 360
Db 301 ILRNPWYCGCKMKWYRDLQSLPVKVNVRGLMQAPEKVRGMAIKDLNAELFCKDSGI 360
QY 361 VSTQITTAIPNTVYPAQOWPAPVTKOPDIKNPKLTKDOQTGSPSRKTTITVKSPTS 420
Db 361 VSTQITTAIPNTVYPAQOWPAPVTKOPDIKNPKLTKDOQTGSPSRKTTITVKSPTS 420
QY 421 DTIHISWKLALPMTALRLSWLKLHSPAFSGITETITVGERSEYLVTALEPDSYKVCW 480
Db 421 DTIHISWKLALPMTALRLSWLKLHSPAFSGITETITVGERSEYLVTALEPDSYKVCW 480
QY 481 PMETSNIYLFDEPVCIEETAPLRMNPPTTLNREOEKPYKNPNPLAAIIGGAVALY 540
Db 481 PMETSNIYLFDEPVCIEETAPLRMNPPTTLNREOEKPYKNPNPLAAIIGGAVALY 540
QY 541 TIALALVCMVVRHNGSLFSRNCAYSKGRKKDDYAEAGTKKDNSLLEIRTSFQMLPIS 600
Db 541 TIALALVCMVVRHNGSLFSRNCAYSKGRKKDDYAEAGTKKDNSLLEIRTSFQMLPIS 600
QY 601 NEPISKEEFVHIHIFPPNGMNLKNNHSESSNSRYSYDSDGIPDSHSHS 649
Db 601 NEPISKEEFVHIHIFPPNGMNLKNNHSESSNSRYSYDSDGIPDSHSHS 649

RESULT 6
US-10-175-737-384
; Sequence 384, Application US/10175737
; Publication No. US20030013153A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430RLC50
; CURRENT APPLICATION NUMBER: US/10/175,737
; CURRENT FILING DATE: 2002-06-19
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 384
; LENGTH: 649
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-175-737-384

```

```

Query Match 100.0%; Score 3400; DB 9; Length 649;
Best Local Similarity 100.0%; Pred. No. 3.3e-261; Indels 0; Gaps 0;
Matches 649; Conservative 0; Mismatches 0;

QY 1 MISAWIFLIGKIGLFLQVAPLSVMAKSCPSVCRDAGFIYCNDRFLTSTPGIPEDA 60
Db 1 MISAWIFLIGKIGLFLQVAPLSVMAKSCPSVCRDAGFIYCNDRFLTSTPGIPEDA 60
QY 61 TTYLQNNQNNAGIPSDKLNLLKVERIYLYHNSLDEFFNLPKYKELHLOENNITIT 120
Db 61 TTYLQNNQNNAGIPSDKLNLLKVERIYLYHNSLDEFFNLPKYKELHLOENNITIT 120
QY 121 YDSLKIPYLEELHLDNDSVSAVSEEGAFRDSNYLLFLSRNHLSTIPWGLPRTTEEL 180
Db 121 YDSLKIPYLEELHLDNDSVSAVSEEGAFRDSNYLLFLSRNHLSTIPWGLPRTTEEL 180
QY 181 RLDNRISTSSSLOGLISLRVLDGNLNNHGLGDKVFFNVLNLTSLVRSNSTAA 240
Db 181 RLDNRISTSSSLOGLISLRVLDGNLNNHGLGDKVFFNVLNLTSLVRSNSTAA 240
QY 241 PVNLPFTNLRLKLYLDNHNRPVPPNAFSLRYLRDMSNNLSNLPQIFDDLDNITQL 300
Db 241 PVNLPFTNLRLKLYLDNHNRPVPPNAFSLRYLRDMSNNLSNLPQIFDDLDNITQL 300
QY 301 ILRNPWYCGCKMKWYRDLQSLPVKVNVRGLMQAPEKVRGMAIKDLNAELFCKDSGI 360
Db 301 ILRNPWYCGCKMKWYRDLQSLPVKVNVRGLMQAPEKVRGMAIKDLNAELFCKDSGI 360
QY 361 VSTQITTAIPNTVYPAQOWPAPVTKOPDIKNPKLTKDOQTGSPSRKTTITVKSPTS 420
Db 361 VSTQITTAIPNTVYPAQOWPAPVTKOPDIKNPKLTKDOQTGSPSRKTTITVKSPTS 420
QY 421 DTIHISWKLALPMTALRLSWLKLHSPAFSGITETITVGERSEYLVTALEPDSYKVCW 480
Db 421 DTIHISWKLALPMTALRLSWLKLHSPAFSGITETITVGERSEYLVTALEPDSYKVCW 480
QY 481 PMETSNIYLFDEPVCIEETAPLRMNPPTTLNREOEKPYKNPNPLAAIIGGAVALY 540
Db 481 PMETSNIYLFDEPVCIEETAPLRMNPPTTLNREOEKPYKNPNPLAAIIGGAVALY 540
QY 541 TIALALVCMVVRHNGSLFSRNCAYSKGRKKDDYAEAGTKKDNSLLEIRTSFQMLPIS 600
Db 541 TIALALVCMVVRHNGSLFSRNCAYSKGRKKDDYAEAGTKKDNSLLEIRTSFQMLPIS 600
QY 601 NEPISKEEFVHIHIFPPNGMNLKNNHSESSNSRYSYDSDGIPDSHSHS 649
Db 601 NEPISKEEFVHIHIFPPNGMNLKNNHSESSNSRYSYDSDGIPDSHSHS 649

RESULT 7
US-10-063-502-132
; Sequence 132, Application US/10063502
; Publication No. US20030023042A1
; GENERAL INFORMATION:
; APPLICANT: Eaton, Dan L.
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3230RLC1
; CURRENT APPLICATION NUMBER: US/10/063,502
; CURRENT FILING DATE: 2002-05-01
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 132
; LENGTH: 649
; TYPE: PRT

```

ORGANISM: Homo Sapien
US-10-063-502-132

Query Match 100.0%; Score 3400; DB 9; Length 649;
Best Local Similarity 100.0%; Pred. No. 3.3e-261;
Matches 649; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MISAAWSIFLTKIGLFLQVAPLSVMAKSCPSVCRDAGFYICNDRLFTSIFGTIPEDA 60
DB 1 MISAAWSIFLTKIGLFLQVAPLSVMAKSCPSVCRDAGFYICNDRLFTSIFGTIPEDA 60
QY 61 TTYLYQNNQINNAGIPSDLKMLKVERIYLYHNSLDEPTNLPKYVKELHLOENNRITIT 120
DB 61 TTYLYQNNQINNAGIPSDLKMLKVERIYLYHNSLDEPTNLPKYVKELHLOENNRITIT 120
QY 121 YDSLSKIPYLEELHDDNSVSAVSIIEGAFRDSNYLRLLFSLRNHLSTIPWGLPRTIEEL 180
DB 121 YDSLSKIPYLEELHDDNSVSAVSIIEGAFRDSNYLRLLFSLRNHLSTIPWGLPRTIEEL 180
QY 181 RLDDNRISTISSPSLQGLTSLKRLVLDGNNLNNHGLGDKVFNVLNLTSLVRSNLTA 240
DB 181 RLDDNRISTISSPSLQGLTSLKRLVLDGNNLNNHGLGDKVFNVLNLTSLVRSNLTA 240
QY 241 PVNLPGTNLRKLYLDQNNHINRVPNAFSLRQLYRLDMSNNLNLPGQIFDLDNITOL 300
DB 241 PVNLPGTNLRKLYLDQNNHINRVPNAFSLRQLYRLDMSNNLNLPGQIFDLDNITOL 300
QY 301 ILRNNPWCGCKMKVROWLQSLPVKNVVRGLMCQAPKVRGMAIKDLNAELFDCXDSGI 360
DB 301 ILRNNPWCGCKMKVROWLQSLPVKNVVRGLMCQAPKVRGMAIKDLNAELFDCXDSGI 360
QY 361 VSTIQTITPAIPNTVYPAQOWPAPVTKQPDKNPKLTQDQOTGSPSRKTTITVKSPTS 420
DB 361 VSTIQTITPAIPNTVYPAQOWPAPVTKQPDKNPKLTQDQOTGSPSRKTTITVKSPTS 420
QY 421 DTHISWKLALPMTALRLSWLKLGHSPAGSITETIVTGERSEYLVTALEPDSPIKVCV 480
DB 421 DTHISWKLALPMTALRLSWLKLGHSPAGSITETIVTGERSEYLVTALEPDSPIKVCV 480
QY 481 PMETSNIYLFDETPVCIEETAPLRMYNPTTLNREQKEPKYKPNPLAALIGGAVALV 540
DB 481 PMETSNIYLFDETPVCIEETAPLRMYNPTTLNREQKEPKYKPNPLAALIGGAVALV 540
QY 541 TIALALVCWYVRNGSLFSRNCAYSGRRKDDYAEAGTKDNSILEIRETSFOMLPTS 600
DB 541 TIALALVCWYVRNGSLFSRNCAYSGRRKDDYAEAGTKDNSILEIRETSFOMLPTS 600
QY 601 NEPTSKKEEFVHTIPPNGMNLKNNHSESSNSRSDSGIPDSHSHS 649
DB 601 NEPTSKKEEFVHTIPPNGMNLKNNHSESSNSRSDSGIPDSHSHS 649

RESULT 8
US-10-173-706-384
Sequence 384, Application US/10173706
Publication No. US2003002293A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C7
CURRENT APPLICATION NUMBER: US/10/173,706
CURRENT FILING DATE: 2002-06-17

Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 612

SEQ ID NO 384
LENGTH: 649
TYPE: PRT
ORGANISM: Homo Sapien
US-10-173-706-384

Query Match 100.0%; Score 3400; DB 9; Length 649;
Best Local Similarity 100.0%; Pred. No. 3.3e-261;
Matches 649; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MISAAWSIFLTKIGLFLQVAPLSVMAKSCPSVCRDAGFYICNDRLFTSIFGTIPEDA 60
DB 1 MISAAWSIFLTKIGLFLQVAPLSVMAKSCPSVCRDAGFYICNDRLFTSIFGTIPEDA 60
QY 61 TTYLYQNNQINNAGIPSDLKMLKVERIYLYHNSLDEPTNLPKYVKELHLOENNRITIT 120
DB 61 TTYLYQNNQINNAGIPSDLKMLKVERIYLYHNSLDEPTNLPKYVKELHLOENNRITIT 120
QY 121 YDSLSKIPYLEELHDDNSVSAVSIIEGAFRDSNYLRLLFSLRNHLSTIPWGLPRTIEEL 180
DB 121 YDSLSKIPYLEELHDDNSVSAVSIIEGAFRDSNYLRLLFSLRNHLSTIPWGLPRTIEEL 180
QY 181 RLDDNRISTISSPSLQGLTSLKRLVLDGNNLNNHGLGDKVFNVLNLTSLVRSNLTA 240
DB 181 RLDDNRISTISSPSLQGLTSLKRLVLDGNNLNNHGLGDKVFNVLNLTSLVRSNLTA 240
QY 241 PVNLPGTNLRKLYLDQNNHINRVPNAFSLRQLYRLDMSNNLNLPGQIFDLDNITOL 300
DB 241 PVNLPGTNLRKLYLDQNNHINRVPNAFSLRQLYRLDMSNNLNLPGQIFDLDNITOL 300
QY 301 ILRNNPWCGCKMKVROWLQSLPVKNVVRGLMCQAPKVRGMAIKDLNAELFDCXDSGI 360
DB 301 ILRNNPWCGCKMKVROWLQSLPVKNVVRGLMCQAPKVRGMAIKDLNAELFDCXDSGI 360
QY 361 VSTIQTITPAIPNTVYPAQOWPAPVTKQPDKNPKLTQDQOTGSPSRKTTITVKSPTS 420
DB 361 VSTIQTITPAIPNTVYPAQOWPAPVTKQPDKNPKLTQDQOTGSPSRKTTITVKSPTS 420
QY 421 DTHISWKLALPMTALRLSWLKLGHSPAGSITETIVTGERSEYLVTALEPDSPIKVCV 480
DB 421 DTHISWKLALPMTALRLSWLKLGHSPAGSITETIVTGERSEYLVTALEPDSPIKVCV 480
QY 481 PMETSNIYLFDETPVCIEETAPLRMYNPTTLNREQKEPKYKPNPLAALIGGAVALV 540
DB 481 PMETSNIYLFDETPVCIEETAPLRMYNPTTLNREQKEPKYKPNPLAALIGGAVALV 540
QY 541 TIALALVCWYVRNGSLFSRNCAYSGRRKDDYAEAGTKDNSILEIRETSFOMLPTS 600
DB 541 TIALALVCWYVRNGSLFSRNCAYSGRRKDDYAEAGTKDNSILEIRETSFOMLPTS 600
QY 601 NEPTSKKEEFVHTIPPNGMNLKNNHSESSNSRSDSGIPDSHSHS 649
DB 601 NEPTSKKEEFVHTIPPNGMNLKNNHSESSNSRSDSGIPDSHSHS 649

RESULT 9
US-10-175-738-384
Sequence 384, Application US/10175738
Publication No. US2003002294A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 ; FILE REFERENCE: P3430R1C45
 ; CURRENT APPLICATION NUMBER: US/10/175,738
 ; CURRENT FILING DATE: 2002-06-19
 ; Prior application removed - See File Wrapper or Palm
 ; NUMBER OF SEQ ID NOS: 612
 ; SEQ ID NO 384
 ; LENGTH: 649
 ; TYPE: PRT
 ; ORGANISM: Homo Sapien
 US-10-175-738-384

Query Match 100.0%; Score 3400; DB 9; Length 649;
 Best Local Similarity 100.0%; Pred. No. 3.3e-261;
 Matches 649; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MISAWSIFLTGKTIGFLQVAPLSVMAKSPSVCRDAGFYCNDRFLTSIPTGIPEDA	60
Db	1	MISAWSIFLTGKTIGFLQVAPLSVMAKSPSVCRDAGFYCNDRFLTSIPTGIPEDA	60
QY	61	TTLYLQNNQINNAGIPSDKLNKLVRIYLYHNSLDEPPTNPKYKELHLOENIRIT	120
Db	61	TTLYLQNNQINNAGIPSDKLNKLVRIYLYHNSLDEPPTNPKYKELHLOENIRIT	120
QY	121	YDSLKIPYLEELHLDNSVSAYSIEEAGFRDSNYLRLLFLSRNHLSTIPWGLPTIEEL	180
Db	121	YDSLKIPYLEELHLDNSVSAYSIEEAGFRDSNYLRLLFLSRNHLSTIPWGLPTIEEL	180
QY	181	RDDNRISTISPSISGLTSLKRLVLDGNLNNHGLGDKVFENLVNLTSLVRSNLTAA	240
Db	181	RDDNRISTISPSISGLTSLKRLVLDGNLNNHGLGDKVFENLVNLTSLVRSNLTAA	240
QY	241	PVNLPTNLRKLYLQDNHINRPPNFAFYRLQYRLDMSNNLSNLPQGFDDLDNITQL	300
Db	241	PVNLPTNLRKLYLQDNHINRPPNFAFYRLQYRLDMSNNLSNLPQGFDDLDNITQL	300
QY	301	ILRNPFWCGCKMWRDNLQSLPVKVVYRGLMCOAPEKVGMAIKDLNAELFDCKDSGI	360
Db	301	ILRNPFWCGCKMWRDNLQSLPVKVVYRGLMCOAPEKVGMAIKDLNAELFDCKDSGI	360
QY	361	VSTIQTITAIPTNVTYPAQGWPAVTKQPDIKNPKLTKDQQTGSPSRKTITITVKSVTS	420
Db	361	VSTIQTITAIPTNVTYPAQGWPAVTKQPDIKNPKLTKDQQTGSPSRKTITITVKSVTS	420
QY	421	DTIHISWKLALPMTALRLSWLKLGHSPAFGSITETIVTGERSEYLVTALEPDPSPYKVCW	480
Db	421	DTIHISWKLALPMTALRLSWLKLGHSPAFGSITETIVTGERSEYLVTALEPDPSPYKVCW	480
QY	481	PMETSNLYLFDETPVCITETAPLRMYNPTTTLNREQKEPKYKPNPLPLAAIIGGAVLV	540
Db	481	PMETSNLYLFDETPVCITETAPLRMYNPTTTLNREQKEPKYKPNPLPLAAIIGGAVLV	540
QY	541	TIALALVGVYVHRNGSLFNRNCAYSKGRRKDDYAEAGTKKNSILEIRETSFQMLPIS	600
Db	541	TIALALVGVYVHRNGSLFNRNCAYSKGRRKDDYAEAGTKKNSILEIRETSFQMLPIS	600
QY	601	NEPISKEEFVHTIFFPPNGMNLKNNHSESSNRSYRSDGIPDSDHSHS	649
Db	601	NEPISKEEFVHTIFFPPNGMNLKNNHSESSNRSYRSDGIPDSDHSHS	649

RESULT 10
 US-10-175-738-384
 ; Sequence 384, Application US/10175752
 ; Publication No. US2003002295A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Chen, Jian
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Gurney, Austin L.

; APPLICANT: Pan, James
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Watanabe, Colin K.
 ; APPLICANT: Wood, William I.
 ; APPLICANT: Zhang, Zemin
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 ; FILE REFERENCE: P3430R1C60
 ; CURRENT APPLICATION NUMBER: US/10/175,752
 ; CURRENT FILING DATE: 2002-06-19
 ; Prior application removed - See File Wrapper or Palm
 ; NUMBER OF SEQ ID NOS: 612
 ; SEQ ID NO 384
 ; LENGTH: 649
 ; TYPE: PRT
 ; ORGANISM: Homo Sapien
 US-10-175-752-384

Query Match 100.0%; Score 3400; DB 9; Length 649;
 Best Local Similarity 100.0%; Pred. No. 3.3e-261;
 Matches 649; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MISAWSIFLTGKTIGFLQVAPLSVMAKSPSVCRDAGFYCNDRFLTSIPTGIPEDA	60
Db	1	MISAWSIFLTGKTIGFLQVAPLSVMAKSPSVCRDAGFYCNDRFLTSIPTGIPEDA	60
QY	61	TTLYLQNNQINNAGIPSDKLNKLVRIYLYHNSLDEPPTNPKYKELHLOENIRIT	120
Db	61	TTLYLQNNQINNAGIPSDKLNKLVRIYLYHNSLDEPPTNPKYKELHLOENIRIT	120
QY	121	YDSLKIPYLEELHLDNSVSAYSIEEAGFRDSNYLRLLFLSRNHLSTIPWGLPTIEEL	180
Db	121	YDSLKIPYLEELHLDNSVSAYSIEEAGFRDSNYLRLLFLSRNHLSTIPWGLPTIEEL	180
QY	181	RDDNRISTISPSISGLTSLKRLVLDGNLNNHGLGDKVFENLVNLTSLVRSNLTAA	240
Db	181	RDDNRISTISPSISGLTSLKRLVLDGNLNNHGLGDKVFENLVNLTSLVRSNLTAA	240
QY	241	PVNLPTNLRKLYLQDNHINRPPNFAFYRLQYRLDMSNNLSNLPQGFDDLDNITQL	300
Db	241	PVNLPTNLRKLYLQDNHINRPPNFAFYRLQYRLDMSNNLSNLPQGFDDLDNITQL	300
QY	301	ILRNPFWCGCKMWRDNLQSLPVKVVYRGLMCOAPEKVGMAIKDLNAELFDCKDSGI	360
Db	301	ILRNPFWCGCKMWRDNLQSLPVKVVYRGLMCOAPEKVGMAIKDLNAELFDCKDSGI	360
QY	361	VSTIQTITAIPTNVTYPAQGWPAVTKQPDIKNPKLTKDQQTGSPSRKTITITVKSVTS	420
Db	361	VSTIQTITAIPTNVTYPAQGWPAVTKQPDIKNPKLTKDQQTGSPSRKTITITVKSVTS	420
QY	421	DTIHISWKLALPMTALRLSWLKLGHSPAFGSITETIVTGERSEYLVTALEPDPSPYKVCW	480
Db	421	DTIHISWKLALPMTALRLSWLKLGHSPAFGSITETIVTGERSEYLVTALEPDPSPYKVCW	480
QY	481	PMETSNLYLFDETPVCITETAPLRMYNPTTTLNREQKEPKYKPNPLPLAAIIGGAVLV	540
Db	481	PMETSNLYLFDETPVCITETAPLRMYNPTTTLNREQKEPKYKPNPLPLAAIIGGAVLV	540
QY	541	TIALALVGVYVHRNGSLFNRNCAYSKGRRKDDYAEAGTKKNSILEIRETSFQMLPIS	600
Db	541	TIALALVGVYVHRNGSLFNRNCAYSKGRRKDDYAEAGTKKNSILEIRETSFQMLPIS	600
QY	601	NEPISKEEFVHTIFFPPNGMNLKNNHSESSNRSYRSDGIPDSDHSHS	649
Db	601	NEPISKEEFVHTIFFPPNGMNLKNNHSESSNRSYRSDGIPDSDHSHS	649

RESULT 11
 US-10-176-482-384
 ; Sequence 384, Application US/10176482
 ; Publication No. US2003002296A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baker, Kevin P.

APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C70
CURRENT FILING DATE: 2002-06-20
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 384
LENGTH: 649
TYPE: PRT
ORGANISM: Homo Sapien
US-10-176-482-384

Query Match 100.0%; Score 3400; DB 9; Length 649;
Best Local Similarity 100.0%; Pred. No. 3.3e-261;
Matches 649; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MISAAWSIFLGTIGTIGLQVAPLSVMAKSCPSVCRDAGFIYCNDRFLTSTPTGIPEDA 60
DB 1 MISAAWSIFLGTIGTIGLQVAPLSVMAKSCPSVCRDAGFIYCNDRFLTSTPTGIPEDA 60
QY 61 TTYLQNNQINNAGIPSDLNKLVKRYIYHNSLDEEPTNPKYKELHLOENIRIT 120
DB 61 TTYLQNNQINNAGIPSDLNKLVKRYIYHNSLDEEPTNPKYKELHLOENIRIT 120
QY 121 YDSLSKIPYLELHLDNDSVSAVSEIEGAFRDSNRYLRLFLSRNHLSTIPWGLPRTIEEL 180
DB 121 YDSLSKIPYLELHLDNDSVSAVSEIEGAFRDSNRYLRLFLSRNHLSTIPWGLPRTIEEL 180
QY 181 RLDDNRISTISSPSLQGLTSLKRLVLDGNLLNHHGLGDKVFNVLNLTSLVRSNTAA 240
DB 181 RLDDNRISTISSPSLQGLTSLKRLVLDGNLLNHHGLGDKVFNVLNLTSLVRSNTAA 240
QY 241 PVNLPCTNLRLKLYLDNHNINRVPNFAFVYLYRQLYRDLMSNNLSNLPQGFDDLDNITQL 300
DB 241 PVNLPCTNLRLKLYLDNHNINRVPNFAFVYLYRQLYRDLMSNNLSNLPQGFDDLDNITQL 300
QY 301 ILRNNPWYCGCKMWVRDLQSLPVKNVVRGLMCOAPEKVRGMAIKDLNAELFDCKDSGI 360
DB 301 ILRNNPWYCGCKMWVRDLQSLPVKNVVRGLMCOAPEKVRGMAIKDLNAELFDCKDSGI 360
QY 361 VSTIQTITTAIPNTVYPAQGWPAFVTKQPDINKPKLTQDQTTGSPSRKTTITTVKSVTS 420
DB 361 VSTIQTITTAIPNTVYPAQGWPAFVTKQPDINKPKLTQDQTTGSPSRKTTITTVKSVTS 420
QY 421 DTTHISWKLALPMTALRLSWLKLGHSPAFGSTETIVTGERSEYLVTALEPDPSPYKVCW 480
DB 421 DTTHISWKLALPMTALRLSWLKLGHSPAFGSTETIVTGERSEYLVTALEPDPSPYKVCW 480
QY 481 PMETSNIYLFDETPVCIETETAPLRMYNPTTINRQEKKEPKNPNLPLAAIIGGAVALV 540
DB 481 PMETSNIYLFDETPVCIETETAPLRMYNPTTINRQEKKEPKNPNLPLAAIIGGAVALV 540
QY 541 TIALALVQWYVHRNGSLFSRNCAYSKGRKDDYAEAGTKKNSILEIRETSFQMLPTS 600
DB 541 TIALALVQWYVHRNGSLFSRNCAYSKGRKDDYAEAGTKKNSILEIRETSFQMLPTS 600
QY 601 NEPISKEEFVIHTIPPPNGMNLKNNHSESSSSNRSDGIPDSDHSHS 649
DB 601 NEPISKEEFVIHTIPPPNGMNLKNNHSESSSSNRSDGIPDSDHSHS 649

RESULT 12

US-10-176-757-384
Sequence 384, Application US/10176757
Publication No. US20030022297A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C86
CURRENT FILING DATE: 2002-06-20
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 384
LENGTH: 649
TYPE: PRT
ORGANISM: Homo Sapien
US-10-176-757-384

Query Match 100.0%; Score 3400; DB 9; Length 649;
Best Local Similarity 100.0%; Pred. No. 3.3e-261;
Matches 649; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MISAAWSIFLGTIGTIGLQVAPLSVMAKSCPSVCRDAGFIYCNDRFLTSTPTGIPEDA 60
DB 1 MISAAWSIFLGTIGTIGLQVAPLSVMAKSCPSVCRDAGFIYCNDRFLTSTPTGIPEDA 60
QY 61 TTYLQNNQINNAGIPSDLNKLVKRYIYHNSLDEEPTNPKYKELHLOENIRIT 120
DB 61 TTYLQNNQINNAGIPSDLNKLVKRYIYHNSLDEEPTNPKYKELHLOENIRIT 120
QY 121 YDSLSKIPYLELHLDNDSVSAVSEIEGAFRDSNRYLRLFLSRNHLSTIPWGLPRTIEEL 180
DB 121 YDSLSKIPYLELHLDNDSVSAVSEIEGAFRDSNRYLRLFLSRNHLSTIPWGLPRTIEEL 180
QY 181 RLDDNRISTISSPSLQGLTSLKRLVLDGNLLNHHGLGDKVFNVLNLTSLVRSNTAA 240
DB 181 RLDDNRISTISSPSLQGLTSLKRLVLDGNLLNHHGLGDKVFNVLNLTSLVRSNTAA 240
QY 241 PVNLPCTNLRLKLYLDNHNINRVPNFAFVYLYRQLYRDLMSNNLSNLPQGFDDLDNITQL 300
DB 241 PVNLPCTNLRLKLYLDNHNINRVPNFAFVYLYRQLYRDLMSNNLSNLPQGFDDLDNITQL 300
QY 301 ILRNNPWYCGCKMWVRDLQSLPVKNVVRGLMCOAPEKVRGMAIKDLNAELFDCKDSGI 360
DB 301 ILRNNPWYCGCKMWVRDLQSLPVKNVVRGLMCOAPEKVRGMAIKDLNAELFDCKDSGI 360
QY 361 VSTIQTITTAIPNTVYPAQGWPAFVTKQPDINKPKLTQDQTTGSPSRKTTITTVKSVTS 420
DB 361 VSTIQTITTAIPNTVYPAQGWPAFVTKQPDINKPKLTQDQTTGSPSRKTTITTVKSVTS 420
QY 421 DTTHISWKLALPMTALRLSWLKLGHSPAFGSTETIVTGERSEYLVTALEPDPSPYKVCW 480
DB 421 DTTHISWKLALPMTALRLSWLKLGHSPAFGSTETIVTGERSEYLVTALEPDPSPYKVCW 480
QY 481 PMETSNIYLFDETPVCIETETAPLRMYNPTTINRQEKKEPKNPNLPLAAIIGGAVALV 540
DB 481 PMETSNIYLFDETPVCIETETAPLRMYNPTTINRQEKKEPKNPNLPLAAIIGGAVALV 540
QY 541 TIALALVQWYVHRNGSLFSRNCAYSKGRKDDYAEAGTKKNSILEIRETSFQMLPTS 600
DB 541 TIALALVQWYVHRNGSLFSRNCAYSKGRKDDYAEAGTKKNSILEIRETSFQMLPTS 600
QY 601 NEPISKEEFVIHTIPPPNGMNLKNNHSESSSSNRSDGIPDSDHSHS 649

Db 601 NEPISKEEFVHTFPPNGMNLKNNHSESSSNRSTDSGIPDSHSHS 649
|||||

RESULT 13

US-10-176-913-384

; Sequence 384, Application US/10176913

; Publication No. US20030022298A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

; APPLICANT: Chen, Jian

; APPLICANT: Desnoyers, Luc

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Pan, James

; APPLICANT: Smith, Victoria

; APPLICANT: Watanabe, Colin K.

; APPLICANT: Wood, William I.

; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

; FILE REFERENCE: P3430RIC66

; CURRENT APPLICATION NUMBER: US/10/176,913

; CURRENT FILING DATE: 2002-06-20

; Prior Application removed - See file Wrapper or Palm

; NUMBER OF SEQ ID NOS: 612

; SEQ ID NO 384

; LENGTH: 649

; TYPE: PRT

; ORGANISM: Homo Sapien

US-10-176-913-384

Query Match 100.0%; Score 3400; DB 9; Length 649;
Best Local Similarity 100.0%; Pred. No. 3.3e-261;
Matches 649; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MISAAWSTFLGKIGLFLQVAPLSVMAKSCPSVCRDAGFYCNDRLFTSIPTGIPEDA 60
|||||
Db 1 MISAAWSTFLGKIGLFLQVAPLSVMAKSCPSVCRDAGFYCNDRLFTSIPTGIPEDA 60
|||||
QY 61 TTYLQNNQINNAGIPSDLKLLKVERIYLYHNSLDEFFNLPKYKELHLOENNITIT 120
|||||
Db 61 TTYLQNNQINNAGIPSDLKLLKVERIYLYHNSLDEFFNLPKYKELHLOENNITIT 120
|||||
QY 121 YDSLKIPYLEELHLDNDSVSAVSTEEGAFRDSNYLRLFLSRNHLSTIPWGLPRTTEEL 180
|||||
Db 121 YDSLKIPYLEELHLDNDSVSAVSTEEGAFRDSNYLRLFLSRNHLSTIPWGLPRTTEEL 180
|||||
QY 181 RDDNRISTSSPSLQGLSLKRLVLDGNLNNHGLGDKVFFNVLNLTSLVRSNSTAA 240
|||||
Db 181 RDDNRISTSSPSLQGLSLKRLVLDGNLNNHGLGDKVFFNVLNLTSLVRSNSTAA 240
|||||
QY 241 PVNLPFTNLRKLYLQDNHINRYPNPAFSYLRQYLRDMSNNLSNLPQGFDDLDNITQL 300
|||||
Db 241 PVNLPFTNLRKLYLQDNHINRYPNPAFSYLRQYLRDMSNNLSNLPQGFDDLDNITQL 300
|||||
QY 301 ILRNNPWTCCGCKMKWVRLNLSLPVKNVIRGLMCOAPEKVRGMAIKDLNAELFCKDSGI 360
|||||
Db 301 ILRNNPWTCCGCKMKWVRLNLSLPVKNVIRGLMCOAPEKVRGMAIKDLNAELFCKDSGI 360
|||||
QY 361 VSTQITTAIPNTVYPAQGWPAVTPKQPDINKPKLTKDQQTGSPSRKTIITIVKSVTS 420
|||||
Db 361 VSTQITTAIPNTVYPAQGWPAVTPKQPDINKPKLTKDQQTGSPSRKTIITIVKSVTS 420
|||||
QY 421 DTIHSWKLALPMTALRLSWLKLGHSPAFGSITETIVTGERSEYLVTALEPDSPPKVCMV 480
|||||
Db 421 DTIHSWKLALPMTALRLSWLKLGHSPAFGSITETIVTGERSEYLVTALEPDSPPKVCMV 480
|||||
QY 481 PMETSNLYLFDTPVCIETETAPLRMYNPTTTLNREQKEPEKYPKNPLPLAAIIGGAVLY 540
|||||
Db 481 PMETSNLYLFDTPVCIETETAPLRMYNPTTTLNREQKEPEKYPKNPLPLAAIIGGAVLY 540
|||||

QY 541 TIALALVCWYVHRNGSLFSRNCAYSKGRRKKDDYAEAGTKKONSILRETSFQMLPIS 600
|||||
Db 541 TIALALVCWYVHRNGSLFSRNCAYSKGRRKKDDYAEAGTKKONSILRETSFQMLPIS 600
|||||
QY 601 NEPISKEEFVHTFPPNGMNLKNNHSESSSNRSTDSGIPDSHSHS 649
|||||
Db 601 NEPISKEEFVHTFPPNGMNLKNNHSESSSNRSTDSGIPDSHSHS 649
|||||

RESULT 14

US-10-180-552-384

; Sequence 384, Application US/10180552

; Publication No. US20030022300A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

; APPLICANT: Chen, Jian

; APPLICANT: Desnoyers, Luc

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Pan, James

; APPLICANT: Smith, Victoria

; APPLICANT: Watanabe, Colin K.

; APPLICANT: Wood, William I.

; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

; FILE REFERENCE: P3430RIC153

; CURRENT APPLICATION NUMBER: US/10/180,552

; CURRENT FILING DATE: 2002-06-25

; Prior Application removed - See File Wrapper or Palm

; NUMBER OF SEQ ID NOS: 612

; SEQ ID NO 384

; LENGTH: 649

; TYPE: PRT

; ORGANISM: Homo Sapien

US-10-180-552-384

Query Match 100.0%; Score 3400; DB 9; Length 649;
Best Local Similarity 100.0%; Pred. No. 3.3e-261;
Matches 649; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MISAAWSTFLGKIGLFLQVAPLSVMAKSCPSVCRDAGFYCNDRLFTSIPTGIPEDA 60
|||||
Db 1 MISAAWSTFLGKIGLFLQVAPLSVMAKSCPSVCRDAGFYCNDRLFTSIPTGIPEDA 60
|||||
QY 61 TTYLQNNQINNAGIPSDLKLLKVERIYLYHNSLDEFFNLPKYKELHLOENNITIT 120
|||||
Db 61 TTYLQNNQINNAGIPSDLKLLKVERIYLYHNSLDEFFNLPKYKELHLOENNITIT 120
|||||
QY 121 YDSLKIPYLEELHLDNDSVSAVSTEEGAFRDSNYLRLFLSRNHLSTIPWGLPRTTEEL 180
|||||
Db 121 YDSLKIPYLEELHLDNDSVSAVSTEEGAFRDSNYLRLFLSRNHLSTIPWGLPRTTEEL 180
|||||
QY 181 RDDNRISTSSPSLQGLSLKRLVLDGNLNNHGLGDKVFFNVLNLTSLVRSNSTAA 240
|||||
Db 181 RDDNRISTSSPSLQGLSLKRLVLDGNLNNHGLGDKVFFNVLNLTSLVRSNSTAA 240
|||||
QY 241 PVNLPFTNLRKLYLQDNHINRYPNPAFSYLRQYLRDMSNNLSNLPQGFDDLDNITQL 300
|||||
Db 241 PVNLPFTNLRKLYLQDNHINRYPNPAFSYLRQYLRDMSNNLSNLPQGFDDLDNITQL 300
|||||
QY 301 ILRNNPWTCCGCKMKWVRLNLSLPVKNVIRGLMCOAPEKVRGMAIKDLNAELFCKDSGI 360
|||||
Db 301 ILRNNPWTCCGCKMKWVRLNLSLPVKNVIRGLMCOAPEKVRGMAIKDLNAELFCKDSGI 360
|||||
QY 361 VSTQITTAIPNTVYPAQGWPAVTPKQPDINKPKLTKDQQTGSPSRKTIITIVKSVTS 420
|||||
Db 361 VSTQITTAIPNTVYPAQGWPAVTPKQPDINKPKLTKDQQTGSPSRKTIITIVKSVTS 420
|||||
QY 421 DTIHSWKLALPMTALRLSWLKLGHSPAFGSITETIVTGERSEYLVTALEPDSPPKVCMV 480
|||||
Db 421 DTIHSWKLALPMTALRLSWLKLGHSPAFGSITETIVTGERSEYLVTALEPDSPPKVCMV 480
|||||

```
QY 481 PMETSNLYLDETPVCIEIETETAPLRMYNPTTILNREQKEPEYKKNPLPLAAIIGGAVALY 540
|||||
Db 481 PMETSNLYLDETPVCIEIETETAPLRMYNPTTILNREQKEPEYKKNPLPLAAIIGGAVALY 540
|||||
QY 541 TIALALVQWYVHRNGSLFSNCAYSKGRRKDDYAEAGTKKDNSILEIRETSFQMLPIS 600
|||||
Db 541 TIALALVQWYVHRNGSLFSNCAYSKGRRKDDYAEAGTKKDNSILEIRETSFQMLPIS 600
|||||
QY 601 NEPIKKEEFVIHTIIPPNGMNLKNNHSESSNRSYRDSGIPDSDHSHS 649
|||||
Db 601 NEPIKKEEFVIHTIIPPNGMNLKNNHSESSNRSYRDSGIPDSDHSHS 649

RESULT 15
US-10-180-557-384
; Sequence 384, Application US/10180557
; Publication No. US20030022301A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Collin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C147
; CURRENT APPLICATION NUMBER: US/10/180,557
; CURRENT FILING DATE: 2002-06-25
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 384
; LENGTH: 649
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-180-557-384

Query Match 100.0%; Score 3400; DB 9; Length 649;
Best Local Similarity 100.0%; Pred. No. 3.3e-261;
Matches 649; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MISAAISFLTGKIGLFLQVAPLSVMAKSCPSVCRCDAGTYCNDRELFISPTGIPEDA 60
|||||
Db 1 MISAAISFLTGKIGLFLQVAPLSVMAKSCPSVCRCDAGTYCNDRELFISPTGIPEDA 60
|||||
QY 61 TTYLQNNQINNAGIPSDLNKLLKVERIYLYHNSLDEFPPTNLPKYKELHLQENNIRITIT 120
|||||
Db 61 TTYLQNNQINNAGIPSDLNKLLKVERIYLYHNSLDEFPPTNLPKYKELHLQENNIRITIT 120
|||||
QY 121 YDSLKIPYLEELHDDNSVSAVIEGAFRDSNYLRLLFLSRNLSLTPWGLPRTIEEL 180
|||||
Db 121 YDSLKIPYLEELHDDNSVSAVIEGAFRDSNYLRLLFLSRNLSLTPWGLPRTIEEL 180
|||||
QY 181 RLDDNRISTSSPSLQGLTSLKRLVLDGNLNNHGLGDKVFNVLNLTSLVRSNLTAA 240
|||||
Db 181 RLDDNRISTSSPSLQGLTSLKRLVLDGNLNNHGLGDKVFNVLNLTSLVRSNLTAA 240
|||||
QY 241 PVNLPGTNLRKLYLDNHNIRVPPNAPSYLRQLYRLDMSNNNLSNLPQIFDDLDNITQL 300
|||||
Db 241 PVNLPGTNLRKLYLDNHNIRVPPNAPSYLRQLYRLDMSNNNLSNLPQIFDDLDNITQL 300
|||||
QY 301 ILRNPWYCGCKMKWVRLQSLPQVKNVRLGMCQAPKVRGMAIKDLNAELFCKDSDGI 360
|||||
Db 301 ILRNPWYCGCKMKWVRLQSLPQVKNVRLGMCQAPKVRGMAIKDLNAELFCKDSDGI 360
|||||
QY 361 VSTIQTITAPNTVYPAQOWPAPVTKOPDIKNPKLTQDQOTGSPSRKTTITTVKSVTS 420
|||||
```

```
Db 361 VSTIQTITAPNTVYPAQOWPAPVTKOPDIKNPKLTQDQOTGSPSRKTTITTVKSVTS 420
|||||
QY 421 DTIHSKWLALPMTALRLSLWKLGHSPAGCSITETVTCERSEYLVTALEPDSPIYKVCWV 480
|||||
Db 421 DTIHSKWLALPMTALRLSLWKLGHSPAGCSITETVTCERSEYLVTALEPDSPIYKVCWV 480
|||||
QY 481 PMETSNLYLDETPVCIEIETETAPLRMYNPTTILNREQKEPEYKKNPLPLAAIIGGAVALY 540
|||||
Db 481 PMETSNLYLDETPVCIEIETETAPLRMYNPTTILNREQKEPEYKKNPLPLAAIIGGAVALY 540
|||||
QY 541 TIALALVQWYVHRNGSLFSNCAYSKGRRKDDYAEAGTKKDNSILEIRETSFQMLPIS 600
|||||
Db 541 TIALALVQWYVHRNGSLFSNCAYSKGRRKDDYAEAGTKKDNSILEIRETSFQMLPIS 600
|||||
QY 601 NEPIKKEEFVIHTIIPPNGMNLKNNHSESSNRSYRDSGIPDSDHSHS 649
|||||
Db 601 NEPIKKEEFVIHTIIPPNGMNLKNNHSESSNRSYRDSGIPDSDHSHS 649
|||||
```

Search completed: April 23, 2003, 11:13:53
Job time : 25 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 11:04:06 ; Search time 21 seconds
(without alignments)
2971.011 Million cell updates/sec

Title: US-10-004-551-6
Perfect score: 3400
Sequence: 1 MISAANSIFLIGTKIGLFLO.....SSNSRYSRDSGIPDSBHS 649

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: PIR_73.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	381.5	11.2	1523	2 J13953	MEGF5 protein - rat
2	376.5	11.1	707	2 J13953	neuronal leucine-r
3	347.5	10.2	1531	2 T42218	slit-1 protein hom
4	344.5	10.1	1489	2 B36665	slit protein 2 pre
5	344.5	10.1	1480	2 A36665	slit protein 1 pre
6	342	10.1	359	1 NBHUC8	decorin precursor
7	342	10.1	360	2 I47020	decorin - rabbit
8	338	9.9	354	2 A55454	decorin precursor
9	336	9.9	360	2 S06280	decorin precursor
10	335.5	9.9	375	2 S55275	fibromodulin precu
11	334	9.8	357	2 S24317	decorin precursor
12	333.5	9.8	375	2 S05390	fibromodulin precu
13	332.5	9.8	354	2 S29145	decorin precursor
14	325	9.6	343	2 A41748	lucican precursor
15	324	9.5	603	2 JC6128	insulin-like growt
16	316.5	9.3	1091	2 A58532	glial cell membran
17	314.5	9.2	605	2 A41915	insulin-like growt
18	311.5	9.2	605	2 JC5239	insulin-like growt
19	309.5	9.1	361	2 A53860	chondroadherin pre
20	306	9.0	603	2 JC1282	insulin-like growt
21	299.5	8.8	380	2 S71876	fibromodulin - chi
22	296.5	8.7	1025	2 T42626	secreted leucine-r
23	285.5	8.4	368	1 BGNHN	biglycan precursor
24	285.5	8.4	369	2 S20811	proteoglycan I - m
25	283.5	8.4	369	2 S32793	biglycan precursor
26	283.5	8.3	369	2 S32559	biglycan precursor
27	273.5	8.0	312	1 NBHUA2	leucine-rich alpha
28	273.5	8.0	338	2 S52284	lumican, secretory
29	270	7.9	382	2 I39068	proline- arginine-

lysine carboxypept
platelet glycoprot
leucine-rich-repea
disease resistance
lumican precursor
orphan G protein-c
oligodendrocyte-my
hypothetical prote
G protein-coupled
platelet membrane
hypothetical prote
hypothetical prote
t1r protein - frui
gene wheeler prote
oncofetal trophobl
oligodendrocyte-my

30 268 7.9 536 2 A34901
31 261.5 7.7 626 1 NBHUA
32 259 7.6 1119 2 AD1822
33 258 7.6 1112 2 T10504
34 257 7.6 342 2 A46743
35 255 7.5 907 2 JE0176
36 249.5 7.3 440 2 A47530
37 249 7.3 680 2 T19939
38 247.5 7.3 907 2 JE0193
39 245.5 7.2 560 2 A60164
40 243 7.1 961 2 T23395
41 242.5 7.1 603 2 T24315
42 240.5 7.1 1385 2 T13887
43 239.5 7.0 1389 2 T13852
44 236 6.9 420 2 A53531
45 235.5 6.9 440 2 A39613

ALIGNMENTS

RESULT 1
T13953
MEGF5 protein - rat
N;Alternate names: slit protein homolog
C;Species: Rattus norvegicus (Norway rat)
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 16-Aug-2002
C;Accession: T13953
R;Nakayama, M.; Nakajima, D.; Nagase, T.; Nomura, N.; Seki, N.; Ohara, O.
Genomics 51, 27-34, 1998
A;Title: Identification of high-molecular-weight proteins with multiple EGF-like mot.
A;Reference number: Z14126; PMID:98360089; PMID:9693030
A;Accession: T13953
A;Status: preliminary; translated from GE/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-1523 <NKA>
A;Cross-references: EMBL:AB011531; NID:g3449291; PIDN:BA32461.1; PID:g3449292
C;Genetics:
C;Superfamily: fruit fly slit protein; EGF homology; leucine-rich alpha-2-glycoprote.

Query Match 11.2%; Score 381.5; DB 2; Length 1523;
Best Local Similarity 22.3%; Pred. No. 1.4e-17;
Matches 179; Conservative 107; Mismatches 272; Indels 243; Gaps 29;

QY 17 LFLQVAPLSVM---AKSCPSVCRCDAAGFYICNDRTFTSIPTGIPEDA----- 60
DB 16 LALALASILSGPPAACPTKCTCSAASVDCHGLGLRAYPRGIPRPAERLDLDNNITR 75
QY 61 -----TTLYLQNNQIN--NAGIPSDLNKLKVERIYLYHNSLDEPPNML---- 102
DB 76 ITKMDFTGLKRLRYLHLEDNQSVIERGAFODLKOL---ERLRNKNKLQVLPFELFQST 132
QY 103 PKYVELHLQENNIITTYDLSLTKPIYLELHLDNSVSASVTEEGAPR----- 151
DB 133 PKLTR-LDLSNQIQIGIPKAFRGVTVGVNQLDNNHISC--IEDGAFRALDLELLTLN 189
QY 152 DSNLYRLFLSRNH----- 165
DB 190 NNNISRLVTSFNHMPKIRTLRLSHNLYCCHLAWLSDWLQRRTTIGQFTLCMAPVHLR 249
QY 166 -----LSTIPWGLPRTI 177
DB 250 GFSVADVQKKYVCPGPHSEAPACNANSLSCPSACSCSNIVDCRGKGLTEIPANLPEGI 309
QY 178 EELRLDDNNRISTISPSLQGLTKRLVLDGLNNGDKVFENLVNLTSLVRSNL 237
DB 310 VEIRLEQNSIKSIAGAFIQYKKLRIDSKNQISD--IAPDAFQGLKSLTSIVLYGNKI 367
QY 238 TAAPVNL-PG-TNLRKLYLDQNNHINVPNAPSVLYRLQYLRDMSNNLSNPQGFIDDL 295
DB 368 TETPKGLFDGLVSLQLLLNNANKINCLRVNTFQDQLNLLISLYDNKLTQISKGLFAPLQ 427

Db 192 NITTPVSSFNHMPKLTFRHLSNHLFCOCHLAWLSQMLRQRTIGLFTQCSGPASLRGL 251
 QY 172 -----GLPRTI 177
 Db 252 NVAEYQKSEFSCGQGEAQAQVACTLSSGSCPAMCSCSNGIVDCRGKGLTAIPANLPETM 311
 QY 178 BEIRLDDNRISTISSPQGLASLKLVLVDGNLNNHGLGDKVFFNLVNLITSLSVNSL 237
 Db 312 TEIRLENGKISPPGAFSPYRKLKIDLSNQIAE--IAPDAFOGRLSGLSVLYGKNI 369
 QY 238 TAAPVNLPG--TNLRKLYLDNHNIRVPPNFAFSLYRLDMSNNNLSNLPQGFDDLD 295
 Db 370 TDLPRGVFGGLTYTLQILLNANKINCIRPDADFODLQNLISLSDYDNKIQSLAKGFTSLR 429
 QY 296 NITQILLNNWYCGCKMKWVRDWLQSLPVKVNVRGLMCAPEKVRGMAIKDLNLELFD 355
 Db 430 AIOHLHAQNFCDCLNKLWDLFLNPIET--GARCASPRILANKRIGQIKSKKPRC 487
 RESULT 4
 B36665
 slit protein 2 precursor - fruit fly (Drosophila melanogaster)
 C:Species: Drosophila melanogaster
 C:Date: 30-Apr-1991 #sequence_revision 30-Apr-1991 #text_change 02-Aug-2002
 C:Accession: B36665
 R:Rothenberg, J.M.; Jacobs, J.R.; Goodman, C.S.; Artavanis-Tsakonas, S.
 Genes Dev. 4, 2169-2187, 1990
 A:Title: slit, an extracellular protein necessary for development of midline glia and co
 A:Reference number: A36665; MUID:91099665; PMID:2176636
 A:Accession: B36665
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-1469 <ROT>
 A:Cross-references: GB:X53959
 C:Genetics:
 A:Gene: FlyBase:slit
 A:Cross-references: FlyBase:FBgn0003425
 C:Superfamily: fruit fly slit protein; EGF homology; leucine-rich alpha-2-glycoprotein
 F:66-91/Domain: proteoglycan amino-terminal homology <PAH1>
 F:101-124/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>
 F:125-148/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>
 F:149-172/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>
 F:173-196/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>
 F:197-220/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>
 F:228-272/Domain: proteoglycan carboxyl-terminal homology <PCS1>
 F:288-313/Domain: proteoglycan amino-terminal homology <PAH2>
 F:323-346/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>
 F:347-370/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR7>
 F:371-394/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR8>
 F:395-418/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR9>
 F:419-442/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR10>
 F:450-494/Domain: proteoglycan carboxyl-terminal homology <PCS2>
 F:512-537/Domain: proteoglycan amino-terminal homology <PAH3>
 F:547-571/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR11>
 F:572-595/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR12>
 F:596-619/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR13>
 F:620-643/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR14>
 F:651-695/Domain: proteoglycan carboxyl-terminal homology <PCS3>
 F:708-733/Domain: proteoglycan amino-terminal homology <PAH4>
 F:743-766/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR15>
 F:767-790/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR16>
 F:846-890/Domain: proteoglycan carboxyl-terminal homology <PCS4>
 F:1028-1061/Domain: EGF homology <EGF>
 F:1068-1099/Domain: EGF homology <EGF2>
 F:1115-1148/Domain: EGF homology <EGF1>
 Query Match 10.18; Score 344.5; DB 2; Length 1469;
 Best Local Similarity 24.08; Pred. No. 4.4e-15;
 Matches 108; Conservative 60; Mismatches 133; Indels 149; Gaps 10;
 QY 22 APLSVMAK--SCPSVCRCDAGFIYCNDRFLTSIPTGTIGEDATLYLQNNAGIPSDLK 80
 Db 285 APMECGAENSCHPCRCADGIVDCREKSLTSVPVTLPPDITDVRLEQN----- 332

QY 81 NLLKVERIYIHNLSDEPFTNLP-----KYKELHLOENNIRITTYDSLSKIPYLEEL 133
 Db 333 -----FTLEPKPSFSSFRRLRIDLNNNLSRIADHSLGKQLTLL 375
 QY 134 HUDONSAYSISIEGAFRDSNYLRLLFLSRNHLSTIPWGLPRTIEE--LRIDDNRISTI 190
 Db 376 VLYGNKIK--DLPSGVFKGLSGSLRLLNANEISCIKDAFRDLHLSLSLYDNNIQL 433
 QY 191 SPSLQGLTSLK----- 202
 Db 434 ANGTFDAMKSNKTVHLAKNPFICDCLNLEWLYLHKNPITETSGARCESPKKMHRRRIESL 493
 QY 203 -----RLVLDGN-----LL 211
 Db 494 REEFKCSWELRKLKSGECRMDSCPMACHEGTVDCTGRRLKEIPRDIPLFTTELL 553
 QY 212 NNHGIG-----DKYFFNLVNLITSLVRSNLSFAFVNI--PGTNLRKLYLDNHNIRVPPN 265
 Db 554 NDNELGRISDGLFGRPLVHLVKELKRNQLGIEPNAFEGASHIQELQIGENKIKESNK 613
 QY 266 AFSYLRQLYRLDMSNNNLSNLPQGFDDLDNITOLIIRNNPWCCKMKWVRDWLQSLPV 325
 Db 614 MFLGLHQLTKTINLYDQISCVMPGSEFHLNLSLTSLNLSNPNCHLAWFAECVRK--K 671
 QY 326 KVNVRGLMCAPEKVRGMAIKDLNLELFD 355
 Db 672 SLNGAARCAGPSKYRQVQIKDLPHSEFKC 701
 RESULT 5
 A36665
 slit protein 1 precursor - fruit fly (Drosophila melanogaster)
 C:Species: Drosophila melanogaster
 C:Date: 30-Apr-1991 #sequence_revision 30-Apr-1991 #text_change 02-Aug-2002
 C:Accession: A36665; A31640; S13523
 R:Rothenberg, J.M.; Jacobs, J.R.; Goodman, C.S.; Artavanis-Tsakonas, S.
 Genes Dev. 4, 2169-2187, 1990
 A:Title: slit, an extracellular protein necessary for development of midline glia an
 A:Reference number: A36665; MUID:91099665; PMID:2176636
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-1480 <ROT>
 A:Cross-references: GB:X53959; NID:g8614; PIDN:CAA37910.1; PID:g8615
 R:Rothenberg, J.M.; Hartley, D.A.; Walther, Z.; Artavanis-Tsakonas, S.
 Cell 55, 1047-1059, 1988
 A:Title: slit; An EGF-homologous locus of D. melanogaster involved in the developmen
 A:Reference number: A31640; MUID:89077533; PMID:3144436
 A:Accession: A31640
 A:Molecule type: DNA
 A:Residues: 881-1182, 'G', 1185-1404, 'GT', 1463-1464, 'YHA', <RO2>
 A:Cross-references: GB:M23543; NID:g340939; PID:g514357
 C:Genetics:
 A:Gene: FlyBase:slit
 A:Cross-references: FlyBase:FBgn0003425
 C:Introns: 1351/3
 C:Superfamily: fruit fly slit protein; EGF homology; leucine-rich alpha-2-glycoprote
 F:66-91/Domain: proteoglycan amino-terminal homology <PAH1>
 F:101-124/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>
 F:125-148/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>
 F:149-172/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>
 F:173-196/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>
 F:197-220/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>
 F:228-272/Domain: proteoglycan carboxyl-terminal homology <PCS1>
 F:288-313/Domain: proteoglycan amino-terminal homology <PAH2>
 F:323-346/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>
 F:347-370/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR7>
 F:371-394/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR8>
 F:395-418/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR9>
 F:419-442/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR10>
 F:450-494/Domain: proteoglycan carboxyl-terminal homology <PCS2>

F:512-537/Domain: proteoglycan amino-terminal homology <PAH3>
 F:547-571/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR11>
 F:572-595/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR12>
 F:596-619/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR13>
 F:620-643/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR14>
 F:651-695/Domain: proteoglycan carboxyl-terminal homology <PCS3>
 F:708-733/Domain: proteoglycan amino-terminal homology <PAH4>
 F:743-766/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR15>
 F:767-790/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR16>
 F:791-814/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR17>
 F:815-838/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR18>
 F:846-890/Domain: proteoglycan carboxyl-terminal homology <PCS4>
 F:1028-1061/Domain: EGF homology <EGF>
 F:1068-1099/Domain: EGF homology <EGF2>
 F:1115-1148/Domain: EGF homology <EGF1>

Query Match 10.1%; Score 344.5; DB 2; Length 1480;
 Best local similarity 24.0%; Pred. No. 4.4e-15;
 Matches 108; Conservative 60; Mismatches 133; Indels 149; Gaps 10;

QY 22 APLSVMAK-SCPSVCRDAGFYCYNDRLSISPTGIPEDATLYLQNNQINNAIPSDLK 80
 Db 285 APMEGAEKSPCHPCADGIVDCKESLSVSPVTLDDTVDVLEQN----- 332
 QY 81 NILKVERIYLYHNSLDPEFTNLP-----KYVKELHQLQNNIRITYDSLSKIPYLEEL 133
 Db 333 -----FTTELPKPSFSSFFRLRRIDLSNNNISRIADHLSGLKQLTTL 375
 QY 134 HLDNVSYSVIEGAFRDSNVLRLFLSRNHLSTIPWGLPTIEE---LRLDNRISTT 190
 Db 376 VLYGNKTK--DLPSGVFGKGLSLRLLLNANEISIRKDAFRLHSLLSLYDNNTQSL 433
 QY 191 SPSLSQGLTSIK----- 202
 Db 434 ANGFDMKSKVTHLAKNPFICDNLWLADYHLKNPITSQARCSFKXMRHRIEEL 493
 QY 203 -----RLVLDGN----- 211
 Db 494 REEFKFSWGLKMKLSGCRMSDCPAMCHGTTVDCTGRRLKEIPRDIPLHTTELL 553
 QY 212 NNHGLG---DKYFENLVNLTSLVRSNLTAAAPVNL--PGTNLRKLYLDQNHINRYPPN 265
 Db 554 NDNEIGRISDGLFGRPLHVLKLEKRNQLGIBPNFAGASHQELQIGENKLIKESNK 613
 QY 266 AFSLYRLQIRLDSNNLSNLPQGFDDLDITOLILRNPNYCCGKMKWRDMLQSLPV 325
 Db 614 MFLGLHQLKTLNLYDQISCVMPGSGFEHLNSLTSNLASPNFNCNHLAWFAECVRK--K 671
 QY 326 KYNVGLMCOAPKVRGMAIKDLNLEFDC 355
 Db 672 SLNGGAARCGAPSKVRDVOIKDLPHEFKC 701

RESULT 6
 NBHUC8
 decorin precursor - human
 N:Alternate names: cartilage proteoglycan protein II; DS-PG II; PG40 core protein; proteoglycan
 C:Species: Homo sapiens (man)
 C:Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 21-Jan-2000
 C:Accession: A45016; A45015; A26476; S05640
 R:Vetter, U.; Vogel, W.; Just, W.; Young, M.F.; Fisher, L.W.
 Genomics 15, 161-168, 1993
 A:Title: Human decorin gene: intron-exon junctions and chromosomal localization.
 A:Reference number: A45016; MUID:93162643; PMID:8432527
 A:Accession: A45016
 A:Molecule type: DNA
 A:Residues: 1-359 <VET>
 A:Cross-references: GB:L01125; GB:L01126; GB:L01127; GB:L01128; GB:L01129; GB:L01130; GB:L01131
 A:Note: sequence extracted from NCBI backbone (NCBIP:125061)
 R:Danielson, K.G.; Razzilo, A.; Cohen, I.; Cannizzaro, L.A.; Eichstetter, I.; Iozzo, R.V.
 Genomics 15, 146-160, 1993
 A:Title: The human decorin gene: intron-exon organization, discovery of two alternative
 A:Reference number: A45015; MUID:93162642; PMID:8432526

A:Accession: A45015
 A>Status: not compared with conceptual translation
 A:Molecule type: DNA
 A:Residues: 38-70 <DA2>
 A:Cross-references: GB:M98262
 A:Note: sequence extracted from NCBI backbone (NCBIP:125013)
 A:Accession: B45015
 A>Status: not compared with conceptual translation
 A:Molecule type: DNA
 A:Residues: 296-359 <DAN>
 A:Note: sequence extracted from NCBI backbone (NCBIP:125017)
 R:Krusius, T.; Ruoslahti, E.
 Proc. Natl. Acad. Sci. U.S.A. 83, 7683-7687, 1986
 A:Title: Primary structure of an extracellular matrix proteoglycan core protein deduced from complementary DNA
 A:Reference number: A26476; MUID:87017013; PMID:3484330
 A:Accession: A26476
 A:Molecule type: mRNA
 A:Residues: 1-359 <RDU>
 A:Cross-references: GB:M14219; NID:gl181169; PIDN:AAB00774.1; PID:gl181170
 R:Roughley, P.J.; White, R.J.
 Biochem. J. 262, 823-827, 1989
 A:Title: Dermatan sulphate proteoglycans of human articular cartilage. The properties of the dermatan sulfate
 A:Reference number: S05639; MUID:90073579; PMID:2590169
 A:Accession: S05640
 A:Molecule type: protein
 A:Residues: 31-33, 'X', 35-50 <ROU>
 C:Comment: This protein binds type I collagen.
 C:Genetics:
 A:Gene: GDB:DCN
 A:Cross-references: GDB:119839; OMIM:125255
 A:Map position: 12q21.3-12q23
 A:Introns: 71/1; 108/3; 180/1; 218/1; 249/2; 295/3
 C:Superfamily: decorin; leucine-rich alpha-2-glycoprotein repeat homology; proteoglycan
 C:Keywords: chondroitin sulfate proteoglycan; collagen binding; dermatan sulfate; dup
 F:1-16/Domain: signal sequence #status predicted <SIG>
 F:17-30/Domain: propeptide #status predicted <PRO>
 F:31-359/Product: decorin #status predicted <MPT>
 F:48-72/Domain: proteoglycan amino-terminal homology <PAH>
 F:82-105/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR11>
 F:106-129/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR12>
 F:130-150/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR13>
 F:151-174/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR14>
 F:175-200/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR15>
 F:201-221/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR16>
 F:222-245/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR17>
 F:246-269/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR18>
 F:270-292/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR19>
 F:293-316/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR20>
 F:308-359/Domain: proteoglycan carboxyl-terminal homology <PCS>
 F:34/Binding site: dermatan sulfate (Ser) (covalent) #status experimental
 F:189,325/Binding site: dermatan sulfate (Ser) (covalent) #status predicted
 F:211,262,303/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 10.1%; Score 342; DB 1; Length 359;
 Best local similarity 31.9%; Pred. No. 8.4e-16;
 Matches 90; Conservative 49; Mismatches 85; Indels 58; Gaps 9;

QY 31 CPSCVCRDAGFYCYNDRLSISPTGIPEDATLYLQNNQINNA--GIPSDLN----- 81
 Db 54 CPFCQCHLVQVCSGLDKVPRDPDPTDITLDLQNNKTEIKDGFKNLHALLIV 113
 QY 82 -----LLKVERIYLYHNSLDPEFTNLPKYVKELHQLQNNI---RTYDLSL 124
 Db 114 NNTSKVSPGAFPTLVKLERLYLSKNOLKELPEKPKTKLQELRAHENEITKVRKVTFGD 173
 QY 125 SKIPLYELHLDNVSYSVIEGAFRDSNVLRLFLSRNHLSTIPWGLPTIEEELRLO 184
 Db 174 NQXIVIE---LGTNPXSGSNGIENGAFQGMKLSYIRIADNTNITSIPQGLPSTELHDG 230
 QY 185 NRITSSPSLOGLTSIKRLVLDGNLLNHLGDKVFFNLVNLTLSELVRSNLTAAAPVNL 244
 Db 231 NKISRVDAASAKLNNNAKLGISFN-----SISAVDNGSLA---NY 268

A:Experimental source: bone
 R:Choi, H.U.; Johnson, T.L.; Pal, S.; Tang, L.H.; Rosenberg, L.; Neame, P.J.
 J. Biol. Chem. 264, 2876-2884, 1989
 A:Title: Characterization of the dermatan sulfate proteoglycans, DS-PGI and DS-PGII, from bovine nasal septum
 A:Reference number: A31430; MUID:89123388; PMID:2914936
 A:Accession: B31430
 A:Molecule type: protein
 A:Residues: 31-33, 'X', 35-54 <CHO>
 A:Experimental source: cartilage; fetal skin
 R:Coster, L.; Rosenberg, L.C.; van der Rest, M.; Poole, A.R.
 J. Biol. Chem. 262, 3809-3812, 1987
 A:Title: The dermatan sulfate proteoglycans of bovine sclera and their relationship to chondroitin sulfate
 A:Reference number: A26545; MUID:87137687; PMID:3818667
 A:Accession: A26545
 A:Molecule type: protein
 A:Residues: 31-50 <COS>
 A:Experimental source: sclera
 R:Pearson, C.H.; Winterbottom, N.; Fackre, D.S.; Scott, P.G.; Carpenter, M.R.
 J. Biol. Chem. 258, 15101-15104, 1983
 A:Reference number: A20935; MUID:84087911; PMID:6654908
 A:Accession: A20935
 A:Molecule type: protein
 A:Residues: 31-54 <PEA>
 A:Experimental source: skin
 R:Chopra, R.K.; Pearson, C.H.; Pringle, G.A.; Fackre, D.S.; Scott, P.G.
 Biochem. J. 232, 277-279, 1985
 A:Title: Dermatan sulphate is located on serine-4 of bovine skin proteodermatan sulphate
 A:Reference number: A44700; MUID:86103195; PMID:3936484
 A:Contents: annotation; glycosylation
 C:Superfamily: decorin; leucine-rich alpha-2-glycoprotein repeat homology; proteoglycan
 C:Keywords: chondroitin sulfate proteoglycan; collagen binding; dermatan sulfate; extracellular matrix
 F:1-15/Domain: signal sequence #status predicted <SIG>
 F:16-30/Domain: propeptide #status predicted <PRO>
 F:31-360/Product: decorin #status predicted <MAT>
 F:49-73/Domain: proteoglycan amino-terminal homology <PAH>
 F:83-106/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>
 F:107-130/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>
 F:131-151/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>
 F:152-175/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>
 F:176-199/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>
 F:202-222/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>
 F:223-246/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR7>
 F:247-270/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR8>
 F:271-293/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR9>
 F:294-308/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR10>
 F:309-360/Domain: proteoglycan carboxyl-terminal homology <PCH>
 F:34/Binding site: dermatan sulfate (Ser) (covalent) #status experimental
 F:190-326/Binding site: dermatan sulfate (Ser) (covalent) #status predicted
 F:212,263,304/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 9.9%; Score 336; DB 2; Length 360;
 Best Local Similarity 28.8%; Pred. No. 2.2e-15;
 Matches 92; Conservative 56; Mismatches 114; Indels 58; Gaps 8;
 QY 18 FQVAPLSVAKSCPSVCRDAGFYCYNDRLTSIPTGIPEDATLYLQNNQINNA--GI 75
 DB 42 FEEVPEHPGVPVPCRCQCHLRVVCSDLGLEKVPKDPDPTDALLDQNNKITEIKD 101
 QY 76 PSDLKN-----LLKVERLYVHNSLDEFTNLPKVKVHELHLEN 114
 DB 102 FKNLKNLHTLLNNKISKISGAPAPLVKLERLYLSKNQKELPEKPKYLOELRVHEN 161
 QY 115 NIRTITVDSKIPYEEELHLDNDSVSAVTEGAFRDSNYLKLFLGRNHLSTIPWGLP 174
 DB 162 EITKVRKSVENGLNQMIVVELGTNPLKSSGTENGAFQMKLSYIRIADTNTITPQLP 221
 QY 175 RIIEELRDDNRISTSSPSLOGLTSKRLVLDGNLLNHLGDKVFFNVLNTELSIVR 234
 DB 222 PSLTEHLHDGNGKTKVDAASLAKLNNLAKLGSF-----SISAVD 262
 QY 235 NSLTAAPVNLPGTNRKLYLDQNHINRVPNNA--FSYLRLYRLDMSNNLSN----- 286

Db 263 NGSLA---NTP--HLRELHNNKLAKVPGVADHYKIQVY---LHNNISALGSNDFC 314
 QY 287 PQGIFDDLONITQLILRNPN 306
 Db 315 PPGYNKKASVGSVLSFNP 334
 RESULT 10
 S55275
 fibromodulin precursor - human
 N:Alternative names: 59K collagen-binding matrix protein
 C:Species: Homo sapiens (man)
 C:Date: 23-Aug-1995 #sequence_revision 03-Oct-1995 #text_change 24-Sep-1999
 A:Accession: S55275; S35710; S32752; S41925
 R:Hildebrand, A.; Komaris, M.; Rasmussen, L.M.; Heinemann, D.; Twardzik, D.R.
 Biochem. J. 302, 527-534, 1994
 A:Title: Interaction of the small interstitial proteoglycans biglycan, decorin and fi
 A:Reference number: S55275; MUID:94379985; PMID:8093006
 A:Accession: S55275
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: mRNA
 A:Residues: 1-376 <HI2>
 A:Cross-references: EMBL:X75546; NID:G453156; PIDN:CAA53233.1; PID:G453157
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1993
 R:Antonsson, P.; Heinemann, D.; Oldberg, A.
 Biochim. Biophys. Acta 1174, 204-206, 1993
 A:Title: Structure and deduced amino acid sequence of the human fibromodulin gene.
 A:Reference number: S35710; MUID:93363641; PMID:8357838
 A:Accession: S35710
 A:Molecule type: DNA
 A:Residues: 1-3, 'T', 5-86, 'P', 88-209, 'N', 211-225, 'Y', 227-343, 'Q', 345-354, 'M', 356-362, 'A'
 A:Cross-references: EMBL:X72913; NID:G297090; PIDN:CAA51418.1; PID:G297091
 C:Comment: Keratan sulfate can be covalently attached to N-acetylglucosamine at the a
 C:Genetics:
 A:Gene: GDB:FM0D
 A:Cross-references: GDB:228979; OMIM:600245
 A:Map position: 1q32.1-1q32.1
 A:Introns: 326/2
 C:Superfamily: fibromodulin; leucine-rich alpha-2-glycoprotein repeat homology
 C:Keywords: cartilage; chondroitin sulfate proteoglycan; collagen binding; connective
 F:1-18/Domain: signal sequence #status predicted <SIG>
 F:19-376/Product: fibromodulin #status predicted <MAT>
 F:38,39,42,45,47,50,53,55,63,65/Binding site: sulfate (Tyr) (covalent) #status predic
 F:127,166,201,291/Binding site: carbohydrate (Asn) (covalent) #status predicted
 Query Match 9.9%; Score 335.5; DB 2; Length 376;
 Best Local Similarity 32.8%; Pred. No. 2.5e-15;
 Matches 97; Conservative 38; Mismatches 116; Indels 45; Gaps 10;
 QY 29 KSCPSVCRDAGFI---YCNDRFTSIPTGIPEDATLYLQNNQI-----NNAG-- 74
 Db 74 RDCPQECDCPPNFLTAMYCDNRNLKYLPPF-VPSRMKYVYFQNNQITSIQGVDFDNATGLL 132
 QY 75 -----IPSD-----LKNLLKVERLYVHNSLDEFTNLPKVKVHELHLENRTIT 120
 Db 133 WIALHGQITSDKVGKVFVKLRLERLYLDHNNLTMPGPPPLRSRLHLDHNSRVP 192
 QY 121 YDSLKIPYEEELHLDNDSVSAVTEGAFRDSNYLKLFLGRNHLSTIPWGLPRTTEL 180
 Db 193 NNALEGLNLTALYLDHDIQEVG---SSMRGLSLLDLSYNHLRKVPDGLFSALEQL 249
 QY 181 RDDNRISTSSPSLOGLTSKRLVLDGNLLNHLGDKVFFNVLNTELSVNSL-FA 239
 Db 250 YMEHNNYITVPDSYTFGAPKLYLVRLSHNSLTNNGLASNT-FNSSSLLELDLSYNLOKI 308
 QY 240 APVNLPGTNRKLYLDQNHINRVPNNAFSL-----RQLYRLDMSNNLSNLP 287
 Db 309 PPVN---TNLENTLQGNKINEFSISSECTVVDVNFVSKLVQVVRDGDGNEIKRSAMP 361
 RESULT 11
 S24317
 decorin precursor - chicken

```

Query Watch          9.8%; Score 333.5; DB 2; Length 375;
Best Local Similarity 32.5%; Pred. No. 3.4e-15;
Matches 96; Conservative 37; Mismatches 119; Indels 43; Gaps 9;

QY 29 KSCPSPVCRCDAGF---IYCNDRFLTPIPGIDEDATTLYLQNNQI-----NNAG-- 74
      : || | | : ||::: | : | : |||| | | | | | | | | | | | | | |
Db 73 RDCPEQCDPPNFPTAMCYCDNRNLKYLFP-VPSRMKYVYFQNNQISSIQGVFDNATGIL 131
      : || | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 75 -----IPSD-----LKNLKLKVERILYINSIDPEPTNLKPYKELHLQENRIRIT 120
      : || | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 132 WIALHGNTSDSKVKKVFKSLRHLERLYLDHNNTTRIPSLPRLSELHLDRHNOISKVP 191
      : || | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 121 YDSLKIPLYLEELHDDNSVSAVIEEGAFROSNTRYLLFLSRNHLSTIPWGIPRTIEEL 180
      : : | : | : | : | : | : | | | | | | | | | | | | | | | | |
Db 192 NNALEGLENTALTYLHNNEIQEVSSMKGLRS---LILDLDSYNHLKRPDGLPSALSQL 248
      : : | : | : | : | : | : | | | | | | | | | | | | | | | | |

```

Biochim. Biophys. Acta 1132, 225-227, 1992
 A:Title: cDNA sequence for rat dermatan sulfate proteoglycan-II (decorin).
 A:Reference number: S29145; MID:93003331; PMID:1390895
 A:Accession: S29145
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-354 <AB>
 A:Cross-references: EMBL:Z12298; NID:g57549; PIDN:CAA78170.1; PID:g57550
 R:Asundi, V.K.; Dreher, K.L.
 Eur. J. Cell Biol. 59, 314-321, 1992
 A:Title: Molecular characterization of vascular smooth muscle decorin: deduced core
 A:Reference number: I60238; MID:93154359; PMID:1493796

Krasinski, V.N.; Dreher, K.L.
 Eur. J. Cell Biol. 59, 314-321, 1992
 A:Title: Molecular characterization of vascular smooth muscle decorin: deduced core
 A:Reference number: I60238; MUID:93154359; PMID:1493796
 A:Accession: I60238
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 11-354 <RCS>
 A:Cross-references: EMBL:X59859; NID:g56056; PID:CAA42519.1; PID:g56057
 C:Genetics:
 A:Gene: DCN
 C:Superfamily: decorin; leucine-rich alpha-2-glycoprotein repeat homology; proteogly
 C:Keywords: collagen binding; extracellular matrix; glycoprotein
 F:1-30/Domain: signal sequence (fragment) #status predicted <SIG>
 F:31-354/Product: decorin #status predicted <MAT>
 F:43-67/Domain: proteoglycan amino-terminal homology <PAH>

Fig. 2 *Neisseria meningitidis* serotype 14 capsular polysaccharide (N14-CP) conjugated with *Neisseria meningitidis* serotype 14 outer membrane vesicles (N14-OMV) and *Neisseria meningitidis* serotype 14 outer membrane vesicles (N14-OMV) conjugated with *Neisseria meningitidis* serotype 14 capsular polysaccharide (N14-CP).

Db 430 LAGISELELDLFRANQTHLPRLFGGLGLEYLELLSNNQLTWLSEDVIGPLQRAFWLDL 489
QY 288 -----QGFDDLDNITQLILN-----NPWYCGCKMKWVRDW 319
:|:| | : | | | | | | | :|:|
Db 490 SHNRLETPAEGFSSGLRFLYLNRNNSIOTFVPPGLERLWLDANPWCSCPLKALRDF 549
QY 320 LQSLPVKY-----NVRGLMCQAPKVRGMAIKDLNLELF 353
:|:| | : | | | | | | | :|:|
Db 550 ALQNEGVVPRFVQVCEGDDCQPVYTYNITCAGPANVSGLDLRDISETLF 600

Search completed: April 23, 2003, 11:06:51
Job time : 26 secs

GenCore version 5.1.4.p5.4578
Copyright (C) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 11:02:51 ; Search time 14 Seconds
(without alignments)
1922.724 Million cell updates/sec

Title: US-10-004-551-6

Perfect score: 3400

Sequence: 1 MISAAWSIFLTKIGLFLQ.....SSNSRYSRDSGIPDSHSHS 649

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3395	99.9	649	1 FLR3_HUMAN	Q9bzu0 homo sapien
2	1920.5	56.5	646	1 FLR1_HUMAN	Q9bnz1 homo sapien
3	1530	45.0	660	1 FLR2_HUMAN	Q4x155 homo sapien
4	350	10.3	360	1 PG22_PIG	Q9xsd9 sus scrofa
5	344.5	10.1	1480	1 SLRT_DROME	P24014 drosophila
6	342	10.1	359	1 PG52_HUMAN	P07585 homo sapien
7	342	10.1	360	1 PG52_RABIT	Q28888 oryctolagus
8	341.5	10.0	376	1 FMOD_HUMAN	Q06828 homo sapien
9	340.5	10.0	376	1 FMOD_MOUSE	P50608 mus musculus
10	340.5	10.0	376	1 FMOD_RAT	P50609 rattus norv
11	338	9.9	354	1 PG52_MOUSE	P28654 mus musculus
12	336	9.9	360	1 PG52_BOVIN	P21793 bos taurus
13	336	9.9	360	1 PG52_SHEEP	Q9tte2 ovis aries
14	334	9.8	356	1 PG52_COTUJA	Q9de68 coturnix co
15	334	9.8	357	1 PG52_CHICK	P28675 gallus gall
16	334	9.8	360	1 PG52_CANFA	Q29393 canis fami
17	333.5	9.8	375	1 FMOD_BOVIN	P13605 bos taurus
18	332.5	9.8	354	1 PG52_RAT	Q01129 rattus norv
19	332.5	9.8	368	1 PG51_XENLA	Q91b75 xenopus lae
20	332	9.8	360	1 PG52_HORSE	Q46542 equus cabal
21	331.5	9.8	379	1 ASPN_HUMAN	Q9bxt1 homo sapien
22	329	9.7	343	1 LDM_COTUJA	Q9de67 coturnix co
23	328	9.6	421	1 OMD_HUMAN	Q99983 homo sapien
24	325	9.6	343	1 LUM_CHICK	P51890 gallus gall
25	324	9.5	603	1 ALS_MOUSE	P70389 mus musculus
26	322.5	9.5	422	1 OMD_BOVIN	Q77742 bos taurus
27	320.5	9.4	373	1 ASPN_MOUSE	Q9bmq4 mus musculus
28	316	9.3	423	1 OMD_RAT	Q9zls7 rattus norv
29	314.5	9.2	605	1 ALS_HUMAN	P35858 homo sapien
30	313.5	9.2	358	1 CHAD_MOUSE	O55226 mus musculus
31	313	9.2	423	1 OMD_MOUSE	Q35103 mus musculus
32	312	9.2	358	1 CHAD_RAT	O70210 rattus norv
33	311.5	9.2	605	1 ALS_PAPHA	O02833 papio hamad

34	308.5	9.1	361	1 CHAD_BOVIN	Q27972 bos taurus
35	306	9.0	603	1 ALS_RAT	P35859 rattus norv
36	299.5	8.8	380	1 FMOD_CHICK	P51887 gallus gall
37	295.5	8.7	713	1 GAC1_HUMAN	O75325 homo sapien
38	293	8.6	359	1 CHAD_HUMAN	O15335 homo sapien
39	285.5	8.4	368	1 PG51_HUMAN	P21810 homo sapien
40	285.5	8.4	369	1 PG51_MOUSE	P28653 mus musculus
41	285.5	8.4	369	1 PG51_RAT	P47853 rattus norv
42	283.5	8.3	369	1 PG51_BOVIN	P21809 bos taurus
43	283.5	8.3	369	1 PG51_SHEEP	O46390 ovis aries
44	283	8.3	966	1 Y918_HUMAN	O94991 homo sapien
45	282.5	8.3	372	1 PG51_HORSE	O46403 equus cabal

ALIGNMENTS

RESULT 1	FLR3_HUMAN	STANDARD;	PRT;	649 AA.
ID	Q9NZU0: Q96KE1; Q9P259; Q96K42; Q96K39;			
DT	15-JUN-2002 (Rel. 41, Created)			
DT	15-JUN-2002 (Rel. 41, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Leucine-rich repeat transmembrane protein FLRT3 precursor			
DE	(Fibronectin-like domain-containing leucine-rich transmembrane protein			
DE	3).			
GN	FLRT3 OR KIRAL469.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A. AND TISSUE SPECIFICITY.			
RX	MEDLINE-20112755; PubMed-10644439;			
RA	Lacy S.E., Bonnemant C.G., Buzney E.A., Kunkel L.M.;			
RT	"Identification of FLRT1, FLRT2, and FLRT3: a novel family of			
RT	transmembrane leucine-rich repeat proteins.";			
RL	Genomics 62:417-426(1999).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Brain;			
RX	MEDLINE-20277482; PubMed-10819331;			
RA	Nagase T., Kikuno R., Ishikawa K.-I., Hirose M., Ohara O.;			
RT	"Prediction of the coding sequences of unidentified human genes. XVII.			
RT	The complete sequences of 100 new cDNA clones from brain which code			
RT	for large proteins in vitro.";			
RL	DNA Res. 7:143-150(2000).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Embryo, and Teratocarcinoma;			
RA	Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,			
RA	Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,			
RA	Wagatsuma M., Hosoi T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,			
RA	Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,			
RA	Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,			
RA	Yamanoto J., Wakamatsu A., Nakamura Y., Nagahara K., Masuho Y.,			
RA	Ninomiya K., Iwayanagi T.;			
RT	"NEDO human cDNA sequencing project.";			
RL	Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE-21638749; PubMed-11780052;			
RA	Deloukas P., Matthews L.R., Ashurst J., Babbage A.K., Bagguley C.L.,			
RA	Jones M., Stavridis G., Almeida J.P., Babbage A.K., Bagguley C.L.,			
RA	Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,			
RA	Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,			
RA	Buck D., Burrill W.D., Butler A.P., Carter C., Carter N.P.,			
RA	Chapman J.C., Ciamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,			
RA	Clegg S., Copley V.E., Collier R.E., Connor R.E., Corby N.R.,			
RA	Coulson A., Coville G.J., Deadman R., Dhani P.D., Dunn M.,			
RA	Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,			
RA	Graham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,			

FT	REPEAT	154	179	199	219	239	259	279	299	319	339	359	379	399	419	439	459	479	499	519	539	559	579	599	619	639	659	679	699	719	739	759	779	799	819	839	859	879	899	919	939	959	979	999	1019	1039	1059	1079	1099	1119	1139	1159	1179	1199	1219	1239	1259	1279	1299	1319	1339	1359	1379	1399	1419	1439	1459	1479	1499	1519	1539	1559	1579	1599	1619	1639	1659	1679	1699	1719	1739	1759	1779	1799	1819	1839	1859	1879	1899	1919	1939	1959	1979	1999	2019	2039	2059	2079	2099	2119	2139	2159	2179	2199	2219	2239	2259	2279	2299	2319	2339	2359	2379	2399	2419	2439	2459	2479	2499	2519	2539	2559	2579	2599	2619	2639	2659	2679	2699	2719	2739	2759	2779	2799	2819	2839	2859	2879	2899	2919	2939	2959	2979	2999	3019	3039	3059	3079	3099	3119	3139	3159	3179	3199	3219	3239	3259	3279	3299	3319	3339	3359	3379	3399	3419	3439	3459	3479	3499	3519	3539	3559	3579	3599	3619	3639	3659	3679	3699	3719	3739	3759	3779	3799	3819	3839	3859	3879	3899	3919	3939	3959	3979	3999	4019	4039	4059	4079	4099	4119	4139	4159	4179	4199	4219	4239	4259	4279	4299	4319	4339	4359	4379	4399	4419	4439	4459	4479	4499	4519	4539	4559	4579	4599	4619	4639	4659	4679	4699	4719	4739	4759	4779	4799	4819	4839	4859	4879	4899	4919	4939	4959	4979	4999	5019	5039	5059	5079	5099	5119	5139	5159	5179	5199	5219	5239	5259	5279	5299	5319	5339	5359	5379	5399	5419	5439	5459	5479	5499	5519	5539	5559	5579	5599	5619	5639	5659	5679	5699	5719	5739	5759	5779	5799	5819	5839	5859	5879	5899	5919	5939	5959	5979	5999	6019	6039	6059	6079	6099	6119	6139	6159	6179	6199	6219	6239	6259	6279	6299	6319	6339	6359	6379	6399	6419	6439	6459	6479	6499	6519	6539	6559	6579	6599	6619	6639	6659	6679	6699	6719	6739	6759	6779	6799	6819	6839	6859	6879	6899	6919	6939	6959	6979	6999	7019	7039	7059	7079	7099	7119	7139	7159	7179	7199	7219	7239	7259	7279	7299	7319	7339	7359	7379	7399	7419	7439	7459	7479	7499	7519	7539	7559	7579	7599	7619	7639	7659	7679	7699	7719	7739	7759	7779	7799	7819	7839	7859	7879	7899	7919	7939	7959	7979	7999	8019	8039	8059	8079	8099	8119	8139	8159	8179	8199	8219	8239	8259	8279	8299	8319	8339	8359	8379	8399	8419	8439	8459	8479	8499	8519	8539	8559	8579	8599	8619	8639	8659	8679	8699	8719	8739	8759	8779	8799	8819	8839	8859	8879	8899	8919	8939	8959	8979	8999	9019	9039	9059	9079	9099	9119	9139	9159	9179	9199	9219	9239	9259	9279
----	--------	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------

QY	65	LQNNQINNAGISPSDKLLKVKRYIYIYHNSLDDEPTNPKYKELHQLQENIRITVDSL	121
Db	70	LHNNQINNAGFAELHNNVQSVHYIYLYGNDDEPMNLKPNVVLHQLQENNIQISRAAL	129
QY	125	SKPILEELHLDNDSVAVSTIEGAFRDSNYLRLLFSRNLHLSITPWGLPTPIELRLD	184
Db	130	AQLAKLEELHLDNDSISTVGVEDGAFRAISLKLELSKNHLSVPGLPYDLQLRVDE	189
QY	195	NRISTISPSLOGLSLAKRLVLDGNLLNHGLGCKVFNFLVNLTELSLVRSNLSFAPVNL	244
Db	190	NRIAVSDMAFONLSERLVTLDGNLLTKNGTAGTTSHTLKLKFSIVRNSLSHPDPL	249
QY	245	PGLNRKLILQDNHINRVPNFAFSYLRQLYRLDMSNNNLNLPGQIFDLDLNIQTILRN	304
Db	250	PGTHLIRLYLDNQNLNHIPLTAFSNLKRLELDISNQLRMLTQGVFDNLSLAKQLTARN	309
QY	305	NPWTGCGKMKWRDMLQSLPVKNVNRGLMCQAPKVRGMAIKDLNLELPCDKSGIVSTI	364
Db	310	NPWFCDCSKWTEWLKYPISLNLVRGFMCOPEQVRGMVRELNNLILSCPTT--TPGL	367
QY	365	QITTAIPNTVTPAQGWAPVTKOP--DIKNP--KLKDKDQOTGTS-----	405
Db	368	PLTFPAPSTA-----SPTQPPILSIPNPSRSYPTPTTKLPTIPDWDGRRVTP	419
QY	406	PSRKHITIVTKSVTSDTHISWKIALPMTALRLSLWLKIGHSFAGSGITETIVTGERSEYL	465
Db	420	PISBRIQLSTHFVNDTSTIQVSWLSLFTVMAYKLTVKMGHSLVGGIVQERIVSGEKQHLS	479
QY	466	VTALEPDSPYKWCMPMTSNLYLFDFTPWCTETAPLRMYNPTTLNRBOKEPKYK--	523
Db	480	LYNLEPSYTRICLPDLDAFNRYADEVTCSEATHASYLNGSNFTASSHEQYTSMSG	538
QY	524	NPNLPLAALIGGAVALVTALLALVCMYVHRNGSLFNRCAYSKGRRRKDDYABAGTKKD	583
Db	539	SPFL--LAGLIGGAVIFVLVLLVTSFCVHMAHKGRTYSOKWKYARG--RRKDDYCYAGTKKD	596
QY	584	NSLLEIETESFQMLPTSNEPISKEEFVHITFPPNGMNLKNNHSESSNSRYSRDSGIPD	643
Db	597	NSLLEMTETSFQIVSLNDQLLKDFLQEIYTPNGGINYTDCH--IPNMRYCNSSVPD	654
QY	644	SDGHS 649	
Db	655	LEHCHT 660	
RESULT 4			
P6S2_PIG			
AC	P6S2_PIG	STANDARD;	PRT; 360 AA.
ID	Q9XSD9; Q9XSH4;		
DT	30-MAY-2000 (Rel. 39, Created)		
DT	30-MAY-2000 (Rel. 39, Last sequence update)		
DT	15-JUN-2002 (Rel. 41, Last annotation update)		
DE	Decorin precursor (Bone proteoglycan II) (P6-52).		
GN	DCN.		
OS	Sus scrofa (Pig).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.		
OC	NCBI_TaxID=9823;		
RP	[1]		
RP	SEQUENCE FROM N.A. (LONG ISOFORM).		
RC	STRAIN=Yorkshire; TISSUE=aorta.		
RA	Stephenson S., Schnoke M., Vesely I.;		
RA	"Cloning of the porcine decorin gene.;"		
RT	Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.		
CC	-!- FUNCTION: May affect the rate of fibrils formation (By		
CC	similarity).		
CC	-!- SUBUNIT: Binds to type I and type II collagen, to fibronectin and		

DR	Pfam; PF00008; EGF; 7.	
DR	Pfam; PF00054; laminin_g; 1.	
DR	Pfam; PF00560; LRR; 16.	
DR	Pfam; PF01462; LRRT; 4.	
DR	Pfam; PF01463; LRRT; 4.	
DR	SMART; SM00041; CT; 1.	
DR	SMART; SM00179; EGF_CA; 2.	
DR	SMART; SM00001; EGF_Like; 5.	
DR	SMART; SM00370; LRR; 4.	
DR	SMART; SM00082; LRRT; 4.	
DR	SMART; SM00013; LRRT; 4.	
DR	SMART; SM00369; LRR_TYP; 9.	
DR	SMART; SM00282; Lemg; 1.	
DR	PROSITE; PS00010; ASX_HYDROXYL; 3.	
DR	PROSITE; PS01185; CTCK_1; 1.	
DR	PROSITE; PS01225; CTCK_2; 1.	
DR	PROSITE; PS00022; EGF_1; 7.	
DR	PROSITE; PS01186; EGF_2; 5.	
DR	PROSITE; PS01187; EGF_CA; 2.	
DR	PROSITE; PS50025; LAM_G_DOMAIN; 1.	
KW	Neurogenesis; Glycoprotein; Signal; Alternative splicing;	
KW	EGF-like domain; Repeat; Leucine-rich repeat.	
FT	SIGNAL 1 36	
FT	CHAIN 37 1480	SLIT PROTEIN.
FT	REPEAT 99 122	LRR 1.
FT	REPEAT 123 146	LRR 2.
FT	REPEAT 148 170	LRR 3.
FT	REPEAT 171 194	LRR 4.
FT	REPEAT 195 218	LRR 5.
FT	REPEAT 220 246	LRR 6.
FT	REPEAT 321 344	LRR 7.
FT	REPEAT 345 368	LRR 8.
FT	REPEAT 369 392	LRR 9.
FT	REPEAT 394 416	LRR 10.
FT	REPEAT 417 440	LRR 11.
FT	REPEAT 522 545	LRR 12.
FT	REPEAT 546 569	LRR 13.
FT	REPEAT 570 593	LRR 14.
FT	REPEAT 595 617	LRR 15.
FT	REPEAT 618 641	LRR 16.
FT	REPEAT 643 666	LRR 17.
FT	REPEAT 678 701	LRR 18.
FT	REPEAT 720 743	LRR 19.
FT	REPEAT 745 764	LRR 20.
FT	REPEAT 765 788	LRR 21.
FT	REPEAT 790 812	LRR 22.
FT	REPEAT 813 836	LRR 23.
FT	REPEAT 838 861	LRR 24.
FT	DOMAIN 907 944	EGF-LIKE 1.
FT	DOMAIN 946 983	EGF-LIKE 2.
FT	DOMAIN 985 1022	EGF-LIKE 3.
FT	DOMAIN 1024 1062	EGF-LIKE 4.
FT	DOMAIN 1064 1100	EGF-LIKE 5.
FT	DOMAIN 1111 1149	EGF-LIKE 6.
FT	DOMAIN 1152 1325	LAMININ G-LIKE.
FT	DOMAIN 1353 1392	EGF-LIKE 7.
FT	DOMAIN 1409 1480	CTCK.
FT	CARBOHYD 111 111	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD 207 207	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD 357 357	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD 435 435	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD 783 783	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD 788 788	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD 958 958	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD 998 998	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD 1060 1060	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD 1159 1159	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD 1175 1175	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD 1243 1243	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD 1292 1292	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD 911 922	BY SIMILARITY.
FT	DISULFID 916 932	BY SIMILARITY.
FT	DISULFID 934 943	BY SIMILARITY.

EMBL:	M14219;	AAB00774.1;	-
EMBL:	L01131;	AAA52301.1;	ALT_SEQ.
EMBL:	L01125;	AAA52301.1;	JOINED.
EMBL:	L01126;	AAA52301.1;	JOINED.
EMBL:	L01127;	AAA52301.1;	JOINED.
EMBL:	L01129;	AAA52301.1;	JOINED.
EMBL:	L01130;	AAA52301.1;	JOINED.
EMBL:	M98262;	AAB60901.1;	-
EMBL:	AF138300;	AAD44713.1;	-
EMBL:	AF138301;	AAF61437.1;	-
EMBL:	AF138302;	AAD44714.1;	-
EMBL:	AF138303;	AAF61438.1;	-
EMBL:	AF138304;	ADA44715.1;	-
EMBL:	EC005322;	AAH05322.1;	-
PIR:	A28476;	NBHUC8.	
PIR:	S05640;	S05640.	
PIR:	B28457;	B28457.	
PIR:	A45016;	A45016.	
Genew:	HGNC;2705;	DCN.	
MM:	125255;	-	
InterPro:	IPRO01611;	LRR.	
InterPro:	IPRO00372;	LRR_Nterm.	
InterPro:	IPRO03592;	LRR_Out.	
InterPro:	IPRO03591;	LRR_TYP.	
Pfam:	PF00560;	LRR; 7.	
Pfam:	PF01462;	LRRNT; 1.	
SMART:	SM00370;	LRR; 3.	
SMART:	SM00113;	LRRNT; 1.	
SMART:	SM00369;	LRR_TYP; 1.	
Glycoptolein:	Extracellular matrix;	Proteoglycan; Repeat;	
KW	Leucine-rich repeat; Signal;	Alternative splicing; Polymorphism.	
FT SIGNAL	1 - 16	POTENTIAL.	
FT PROPEP	17 - 30		
FT CHAIN	31 - 359	DECORIN.	
FT DOMAIN	54 - 67	CYS-RICH.	
FT REPEAT	94 - 117	LRR-S 1.	
FT REPEAT	118 - 141	LRR-T 1.	
FT REPEAT	142 - 162	LRR-T 2.	
FT REPEAT	163 - 186	LRR-S 2.	
FT REPEAT	187 - 212	LRR-T 3.	
FT REPEAT	213 - 233	LRR-T 4.	
FT REPEAT	234 - 257	LRR-S 3.	
FT REPEAT	258 - 281	LRR-T 5.	
FT REPEAT	282 - 304	LRR-T 6.	
FT REPEAT	305 - 334	LRR-S 4.	
FT REPEAT	335 - 359	LRR-T 7.	
FT DISULEFD	54 - 67	LRR-T 8.	
FT DISULEFD	313 - 346	BY SIMILARITY.	
FT CARBOHYD	34 - 34	O-LINKED (GLYCOSAMINOGLYCAN).	
FT CARBOHYD	211 - 211	N-LINKED (GLCNAC. .) (POTENTIAL).	
FT CARBOHYD	262 - 262	N-LINKED (GLCNAC. .) (POTENTIAL).	
FT CARBOHYD	303 - 303	N-LINKED (GLCNAC. .) (POTENTIAL).	
FT VASPLIC	71 - 179	MISSING (IN ISOFORM B).	
FT VASPLIC	73 - 219	MISSING (IN ISOFORM C).	
FT VASPLIC	109 - 295	MISSING (IN ISOFORM D).	
FT VASPLIC	72 - 75	LKV -> CLPS (IN ISOFORM E).	
FT VASPLIC	76 - 359	MISSING (IN ISOFORM E).	
FT VARIANT	273 - 273	E -> Q (IN DBSNF:1803344).	
CONFLICT	37 - 37	/FTIG-Var_011975.	
CONFLICT	45 - 45	G -> A (IN REF. 7).	
CONFLICT	359 AA;	D -> P (IN REF. 7).	
SEQUENCE	359 AA;	FF51E871A1A5ADD CRC64;	
Query Match	10.1%;	Score 342; DB 1; Length 359;	
Best Local Similarity	31.9%;	Pred. No. 7.6e-16;	
Matches	90; Conservative	49; Mismatches 85; Indels 58; Gaps	
QY	31 CPVCRCDAGFYCHDRFLSIPTGTPEATFLYLQNQFNNA--GI PSDLKN-----	81	
DB	: :		
DB	54 CPFRQCQHVRCQSDLGIDKVPDFDPDTTLLDLQNNKITKDGDFKNLKLHLIV 1134		
OV	82 -----LLKVERIYYHNSDEFFTNLPKYVELHLOENNI---RTTYDSL 124		

```

DB 114 NKLKSVPGAFPLVLRKLYLGNKOLKELPEKPKTLQELRAHENEITKVRKVTENG 173
QY 125 SKIPLLELHLDNSVSAVTEGAFRDSNYLRLLFLSRNHLSTIPWGLPRTTEELRLDD 184
DB 174 NQIVIB--LGVNPLKSSGIENGAFQGMKKLSYIRIADTNITSPQGLPSPTEHLIDG 230
QY 185 NRISTSPSLOGITSLKRLVLDGNLNNHGLGDKVFFNLVNLTLFSLVRSNLTAAAPVNL 244
DB 231 NKISVDAASLGNLNLAKLGSFN-----SISAVNGSLA---NF 268
QY 245 PGTNRKLYLDNHNIRVPPN--AFSYRLQYRLDMSNNLS 284
DB 269 P--HLRELHLDNNKLTRVPGGLAEHKYIQVY---LHNNNIS 305

RESULT 7
PGS2_RABIT
ID QGS2_RABIT STANDARD; PRT; 360 AA.
AC Q28888; Q28608;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Decorin precursor (Bone Proteoglycan II) (PG-S2).
GN DCM.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
SEQUENCE FROM N.A.
RC TISSUE=Cornea;
RX MEDLINE=95122319; PubMed=7822148;
RA Zhan Q., Burrows R., Cintron C.;
RT "Cloning and in situ hybridization of rabbit decorin in corneal
RL tissues."
RL Invest. Ophthalmol. Vis. Sci. 36:206-215(1995).
RN [2]
SEQUENCE OF 38-358 FROM N.A.
RC TISSUE=Cartilage;
RA Hering T.M., Kollar J.;
RT "The primary structure of rabbit chondrocyte decorin deduced from
RL nucleotide sequence."
RL Submitted (NOV-1993) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: May affect the rate of fibrils formation (By
CC similarity).
CC -!- SUBUNIT: Binds to type I and type II collagen, to fibronectin and
CC TGF-beta. Forms a ternary complex with MFAP2 and ELN (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Extracellular matrix (By similarity).
CC -!- PM: The attached glycosaminoglycan chain can be either
CC chondroitin sulfate or dermatan sulfate depending upon the tissue
CC of origin (By similarity).
CC -!- SIMILARITY: BELONGS TO THE SMALL LEUCINE-RICH PROTEOGLYCAN
CC (SLRPS) FAMILY. CLASS I SUBFAMILY.
CC -!- SIMILARITY: CONTAINS 12 LEUCINE-RICH REPEATS (LRR).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL; S76584; AAC33083.1; -
CC EMBL; U03394; AAC04315.1; -
CC InterPro; IPR001611; LRR
CC InterPro; IPR00372; LRR_Nterm.
CC InterPro; IPR003592; LRR_out.
CC InterPro; IPR003591; LRR_Typ.
CC Pfam; PF00560; LRR; 7.
CC Pfam; PF01462; LRRNT; 1.

```

```

DR SMART; SM00370; LRR; 1.
DR SMART; SM00013; LRRNT; 1.
DR SMART; SM00369; LRR_Typ; 1.
KW Glycoprotein; Extracellular matrix; Proteoglycan; Repeat;
KW Leucine-rich repeat; Signal.
FT SIGNAL 1 16
FT PROPEP 17 30
FT CHAIN 31 360
FT DOMAIN 55 68
FT REPEAT 74 94
FT REPEAT 95 118
FT REPEAT 119 142
FT REPEAT 143 163
FT REPEAT 164 187
FT REPEAT 188 213
FT REPEAT 214 234
FT REPEAT 235 258
FT REPEAT 259 282
FT REPEAT 283 305
FT REPEAT 306 335
FT REPEAT 336 360
FT DISULFID 55 68
FT DISULFID 314 347
FT CARBOHYD 34 34
FT CARBOHYD 212 212
FT CARBOHYD 263 263
FT CARBOHYD 304 304
SQ SEQUENCE 360 AA; 39896 MW; 0B50C675FE02369 CRC64;

Query Match 10.1%; Score 342; DB 1; Length 360;
Best Local Similarity 30.9%; Pred. No. 7.6e-16;
Matches 90; Conservative 50; Mismatches 97; Indels 54; Gaps 7;

QY 20 QVAPLSVMAKSPSCRCDAAGFYCNDREFTSTPGIPEDATFTYLLQNNQINNA--GIPS 77
DB 44 ELPDLMLGVPVCFRCQCHLRVQCSDLGLDKVPKDLPPDFTLLDQNNKITKDGDFK 103
QY 78 DLKN-----ILKVERLYVHNSIDFPTNPKYVKELHQENNI 116
DB 104 NLKMLHALIIVNNKISKISPGAFPLVLRKLYLGNKOLKELPEKPKSLQELRAHENEI 163
QY 117 RTITYDSLKIPYLEELHLDNSVSAVTEGAFRDSNYLRLLFLSRNHLSTIPWGLPRT 176
DB 164 TKYKSVFSQNMQIVIELGTNPLKSSGIENGAFQGMKKLSYIRIADTNITTTIPQGLP 223
QY 177 IEELRLDNNRISTSPSLOGITSLKRLVLDGNLNNHGLGDKVFFNLVNLTLFSLVRN- 235
DB 224 LTELHLDGNNKLYIKIDASSKGLNKLAKLGSFN-----DISAVNG 264
QY 236 SLTAAPVNLPGTNRKLYLDNHNIRVPPN--AFSYRLQYRLDMSNNLS 284
DB 265 SLANAP-----HLRELHLDNNKLIRVPGGLADHKYIQVY---LHNNNIS 306

RESULT 8
FMOD_HUMAN
ID FMOD_HUMAN STANDARD; PRT; 376 AA.
AC Q06828; Q15331;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Fibromodulin precursor (FM) (Collagen-binding 59 kDa protein) (Keratan
DE sulfate proteoglycan fibromodulin) (KSPG fibromodulin).
GN FMOD OR FM.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE=93363641; PubMed=8357838;
RA Antonsson P., Heinegaard D., Oldberg A.;

```

RT "Structure and deduced amino acid sequence of the human fibromodulin
 RL gene.";
 RL Biochim. Biophys. Acta 1174:204-206(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Hildebrand A., Twardzik G., Border W.A., Ruoslahti E.;
 RL Submitted (FEB-1994) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: Affects the rate of fibrils formation. May have a
 CC primary role in collagen fibrillogenesis (By similarity).
 CC -!- SUBUNIT: Binds to type I and type II collagen (By similarity).
 CC -!- SUBCELLULAR LOCATION: Extracellular matrix.
 CC -!- PTM: Binds keratan sulfate chains.
 CC -!- SIMILARITY: BELONGS TO THE SMALL LEUCINE-RICH PROTEOGLYCANS
 CC (SLRPS) FAMILY. CLASS II SUBFAMILY.
 CC -!- SIMILARITY: CONTAINS 12 LEUCINE-RICH REPEATS (LRR).
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; X72913; CAA51418.1; -;
 DR EMBL; X75546; CAA53233.1; -;
 DR PIR; S32752; S32752.
 DR PIR; S35710; S35710.
 DR Genew; HGNC:3774; FMOD.
 DR MIM; 600245; -;
 DR InterPro; IPR001611; LRR.
 DR InterPro; IPR000372; LRR_Nterm.
 DR InterPro; IPR003592; LRR_out.
 DR InterPro; IPR003591; LRR_typ.
 DR Pfam; PF00560; LRR; 8.
 DR Pfam; PF01462; LRRNT; 1.
 DR SMART; SM00370; LRR; 1.
 DR SMART; SM00013; LRRNT; 1.
 DR SMART; SM00369; LRR_Typ; 1.
 DR Glycoprotein; Extracellular matrix; Proteoglycan; Repeat;
 KW leucine-rich repeat; signal.
 FT SIGNAL 1 18 BY SIMILARITY.
 FT CHAIN 19 376 FIBROMODULIN.
 FT DOMAIN 76 92
 FT REPEAT 98 117 LRR-S 1.
 FT REPEAT 118 141 LRR-T 1.
 FT REPEAT 142 167 LRR-T 2.
 FT REPEAT 168 188 LRR-S 2.
 FT REPEAT 189 212 LRR-T 3.
 FT REPEAT 213 235 LRR-T 4.
 FT REPEAT 236 256 LRR-S 3.
 FT REPEAT 257 280 LRR-T 5.
 FT REPEAT 281 305 LRR-T 6.
 FT REPEAT 306 325 LRR-S 4.
 FT REPEAT 326 355 LRR-T 7.
 FT REPEAT 356 376 LRR-T 8.
 FT DISULFID 334 367
 FT CARBOHYD 127 127
 FT CARBOHYD 166 166
 FT CARBOHYD 201 201
 FT CARBOHYD 291 291
 FT CARBOHYD 341 341
 FT CONFLICT 4 4
 FT CONFLICT 87 87
 FT CONFLICT 210 210
 FT CONFLICT 226 226
 FT CONFLICT 344 344
 FT CONFLICT 355 355
 FT CONFLICT 363 363

SQ SEQUENCE 376 AA; 43246 MW; C2A78E0678EC3E72 CRC64;
 Query Match 10.0%; Score 341.5; DB 1; Length 376;
 Best Local Similarity 33.1%; Pred. No. 8.8e-16;
 Matches 98; Conservative 37; Mismatches 116; Indels 45; Gaps 10;
 QY 29 KSCPSVCRCDAGF---LYCNDREFTSPTGTPEDATNLYLQNNQI-----NNAG-- 74
 Db 74 RDCPQECDCPPNPFTAYCDNNKLYLFF-VPSRMKYVTFQNNQITQSGVDFNATGLL 132
 QY 75 -----IPSD-----LKNLLKVERIYLYHNSLDEFTPLNPKYVKELHLOENIRIT 120
 Db 133 WIALHGNOITSDKVGKVFESKLRLHLRLYLDDHNNLTMPGCLPRSLRELHLDHNOISRPV 192
 QY 121 YDSLSKPYLEELHDDNSVSVSIEGAFRDSNYLRLLFLSRNHLSTIPWGLPRTIEEL 180
 Db 193 NNALEGENITALLYLQHNIEQVGV---SSMRGLRSLYLLDLSYNHLRKPVDGLPSALEQL 249
 QY 181 RLDNRITSTSSPSLQGLTSKRLVLDGNNLHGLGDKVFENLVNLTFLSLVNSL-TA 239
 Db 250 YMEHNNYVTPDSYFRGAPKLLYLVLSHNSLTNNGLASNT-FNSSLLLELDLSYNQLOKI 308
 QY 240 APVNLPGTNLRKLYLODHNINRVPNAPSYL-----RQLYRLDMNNNLSNP 287
 Db 309 PPVN---TNLENLYLQGNRINEFSISFCTVVDVWVNSQLQVVRDGNEMKRSAMP 361
 RESULT 9
 FMOD_MOUSE STANDARD; PRT; 376 AA.
 AC P50608;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Fibromodulin precursor (PM) (Collagen-binding 50 kDa protein) (Keratan
 sulfate proteoglycan fibromodulin) (KSPG fibromodulin).
 GN FMOD.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Cartilage;
 RX PubMed=11311118;
 RA Seamaenen A.-M.K., Salminen H.J., Rantakokko A.J., Heinigaard D.,
 Vuorio E.I.;
 RT "Murine fibromodulin: cDNA and genomic structure, and age-related
 expression and distribution in the knee joint.";
 RL Biochem. J. 355:577-585(2001).
 CC -!- FUNCTION: Affects the rate of fibrils formation. May have a
 CC primary role in collagen fibrillogenesis.
 CC -!- SUBUNIT: Binds to type I and type II collagen.
 CC -!- SUBCELLULAR LOCATION: Extracellular matrix.
 CC -!- TISSUE SPECIFICITY: Highest levels observed in knee epiphysis, in
 CC calvarial and diaphyseal bone, in nasal and costal cartilage, in
 CC the eye, and in bladder. In mature knee joint it is mostly present
 CC in the proliferating zone of growth plate. It is also observed in
 CC ligaments, especially at insertion sites, in the junction between
 CC meniscus and joint capsule, in the perimysium of skeletal muscle
 CC and in the periosteum.
 CC -!- DEVELOPMENTAL STAGE: Highest levels between 5 days and 1 month of
 CC age. Thereafter, the expression of declined to a level of approx.
 CC 35% of maximum, and remained constant throughout the rest of the
 CC observation period.
 CC -!- PTM: Binds keratan sulfate chains (By similarity).
 CC -!- PTM: Sulfated on tyrosine residue(s) (Probable).
 CC -!- SIMILARITY: BELONGS TO THE SMALL LEUCINE-RICH PROTEOGLYCANS
 CC (SLRPS) FAMILY. CLASS II SUBFAMILY.
 CC -!- SIMILARITY: CONTAINS 12 LEUCINE-RICH REPEATS (LRR).
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; X94998; CAA64454.1; --
 DR MGD; MGI:1328364; Fmod.
 DR InterPro; IPR001611; LRR.
 DR InterPro; IPR000372; LRR_NTerm.
 DR InterPro; IPR003592; LRR_Out.
 DR InterPro; IPR003591; LRR_Typ.
 DR Pfam; PF00560; LRR; 9.
 DR Pfam; PF01462; LRRNT; 1.
 DR SMART; SM00370; LRR; 1.
 DR SMART; SM00013; LRRNT; 1.
 DR SMART; SM00369; LRR_TYP; 1.
 DR Glycoprotein; Extracellular matrix; Sulfation.
 KW Leucine-rich repeat; Signal; Sulfate.
 FT SIGNAL 1 18
 FT CHAIN 19 376
 FT DOMAIN 76 92
 FT REPEAT 98 117
 FT REPEAT 118 141
 FT REPEAT 142 167
 FT REPEAT 168 188
 FT REPEAT 189 212
 FT REPEAT 213 235
 FT REPEAT 236 256
 FT REPEAT 257 280
 FT REPEAT 306 325
 FT REPEAT 326 355
 FT REPEAT 356 376
 FT DOMAIN 68 73
 FT DISULFID 334 367
 FT CARBOHYD 127 127
 FT CARBOHYD 166 166
 FT CARBOHYD 201 201
 FT CARBOHYD 201 201
 FT CARBOHYD 291 291
 FT CARBOHYD 341 341
 FT CARBOHYD 376 AA; 43054 MW; 45A9EDDE0BADA85B CRC64;

Query Match 10.0%; Score 340.5; DB 1; Length 376;
 Best Local Similarity 33.3%; Pred. No. 1e-15;
 Matches 101; Conservative 36; Mismatches 121; Indels 45; Gaps 10;

QY 22 APLSVMAKSCPSVRCDAAGF---IYCNDRFLTSIPTGIPEDATLYLONNOI----- 70
 DB 67 APPPEPRDCPCQECDCFPNPTAMCDNRNLKLPF-VPSRMKYVYFQNNQISAIQGVF 125
 QY 71 -NNAG-----TPSD-----LKNLLAVRILYHNSLDEPTNLPKVVKELHQE 113
 DB 126 DNATGLUWALHGNQITSDVKGVKFSKLRHLRLYLDHNNTRMPGPPSLRELHLDH 185
 QY 114 NNIRITTYDSLSKIPYLEELHLDNDSVSAVIEGAFRDSNRYLRLFLSRNHLSTIPWGL 173
 DB 186 NOISRVPNNALEGENLTALYLHNEIQEVG---SSMRGLRSLLLDLSYNHLRVPDGL 242
 QY 174 PTFIEELDRNRITISISSPSQGLTSLKRLVLDGNNLNHGHGDKVFNVLNTELSLV 233
 DB 243 PSALQLELHNNTYTPDSYRGSFKLLYRLSHNSLTNNGLATNT-FNSSSLFELDL 301
 QY 234 RNSL-TAAPVNLPTNLRKLYLDHNRVPPNAFSL-----ROLRLDMNNNLS 284
 DB 302 YNLOKIPPNV---TNLENYLQGNHNEFSISFCTVDVDMFVKLQVRLDGNKIRS 358
 QY 285 NLP 287
 :!

Db 359 AMP 361
 RESULT 10
 FMOD_RAT FMOD_RAT STANDARD; PRT; 376 AA.
 AC PS0609;
 DT 01-OCT-1996 (Rel. 34, last sequence update)
 DT 01-OCT-1996 (Rel. 34, last sequence update)
 DT 15-JUN-2002 (Rel. 41, last annotation update)
 DE Fibromodulin precursor (FW) (Collagen-binding 59 kDa protein) (Keratan
 DE sulfate proteoglycan fibromodulin) (KSPG fibromodulin).
 GN FMOD.
 OS Rattus norvegicus (Rat).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley;
 RA Krull N.;
 RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: Affects the rate of fibrils formation. May have a
 CC primary role in collagen fibrillogenesis (By similarity).
 CC -!- SUBUNIT: Binds to type I and type II collagen (By similarity).
 CC -!- SUBCELLULAR LOCATION: Extracellular matrix.
 CC -!- PTM: Binds keratan sulfate chains (By similarity).
 CC -!- SIMILARITY: BELONGS TO THE SMALL LEUCINE-RICH PROTEOGLYCAN
 CC (SLRPS) FAMILY. CLASS II SUBFAMILY.
 CC -!- SIMILARITY: CONTAINS 12 LEUCINE-RICH REPEATS (LRR).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; X82152; CAA57648.1; --
 DR InterPro; IPR001611; LRR.
 DR InterPro; IPR000372; LRR_NTerm.
 DR InterPro; IPR003592; LRR_Out.
 DR InterPro; IPR003591; LRR_Typ.
 DR Pfam; PF00560; LRR; 9.
 DR Pfam; PF01462; LRRNT; 1.
 DR SMART; SM00370; LRR; 1.
 DR SMART; SM00013; LRRNT; 1.
 DR SMART; SM00369; LRR_TYP; 1.
 KW Glycoprotein; Extracellular matrix; Proteoglycan; Repeat;
 KW Leucine-rich repeat; Signal.
 FT SIGNAL 1 18
 FT CHAIN 19 376
 FT DOMAIN 76 92
 FT REPEAT 98 117
 FT REPEAT 118 141
 FT REPEAT 142 167
 FT REPEAT 168 188
 FT REPEAT 189 212
 FT REPEAT 213 235
 FT REPEAT 236 256
 FT REPEAT 257 280
 FT REPEAT 281 305
 FT REPEAT 306 325
 FT REPEAT 326 355
 FT REPEAT 356 376
 FT DOMAIN 68 73
 FT DISULFID 334 367
 FT CARBOHYD 127 127
 FT CARBOHYD 166 166
 FT CARBOHYD 201 201
 FT CARBOHYD 201 201
 FT CARBOHYD 291 291
 FT CARBOHYD 341 341
 FT CARBOHYD 376 AA; 43054 MW; 45A9EDDE0BADA85B CRC64;

```

FT CARBOHYD 291 291 (BY SIMILARITY).
FT N-LINKED (GLCNAC. . .) (KERATAN SULFATE)
FT CARBOHYD 341 341 (BY SIMILARITY).
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 376 AA; 43219 MW; 9C3298675CE3714A CRC64;

Query Match
Best Local Similarity 33.3%; Pred. No. 1e-15;
Matches 101; Conservative 36; Mismatches 121; Indels 45; Gaps 10;

QY 22 APLSVMAKSPVCRDAGF---IYCDNRLFTSIPTGIPEDATHTYLNQNI----- 70
Db 67 APPPEPRDCQECDCPPNFTAMCDNRNLKYLPP-VPSRMKYVYFQNNQIAALQGVF 125
QY 71 NNAG-----IPSD-----LKNLLKVERIYLYHNSLDEPTNPKYVRELHQE 113
Db 126 DNATGLLWALHNGQTSKIGKVFSLRHLRLYLDHNNLFRMPGUPRSLRELHLH 185
QY 114 NNIRTYDSKIPYLEELDDNSVSAVSEIEGAFRDSNYLRLLFLSRNHLSTIPWGL 173
Db 186 NQISRPVNNALGLENTALYLHNEIQEVG---SSMRGLRSLILLDLSYNHLRRVDPGL 242
QY 174 PTIEELRLDNRLSISPSQGTSTSLKRLVLDGNNLNNHGLGDKVFNLYNLTLSIV 233
Db 243 PSALQELVLENNYVTPDSYFSGSKLLLYKRLSHNSLTNNGLAUNT-FNSSLIELDLS 301
QY 234 RNSL-TAAPVNLPGTNRKLYLDHNNKINRVPNFAFSYL-----RQYRLDMSNNLS 284
Db 302 YNQLQKIPPNV---TNLENLYLQGNRINEFSISFCVVDVYMFSLQVRLDGNKRS 358
QY 285 NLP 287
Db 359 AMP 361

RESULT 11
PGS2_MOUSE STANDARD; PRT; 354 AA.
AC P28654;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Decorin precursor (Bone proteoglycan II) (PG-S2) (PG40).
GN DCN.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NIH Swiss; TISSUE=Fibroblast;
RA Naitoh Y., Suzuki S.;
RL Submitted (JUL-1990) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=95050610; PubMed=7961765;
RA Schoizen T., Solursch M., Suzuki S., Reiter R., Morgan J.L.,
RA Buchberg A.M., Siracusa L.D., Iozzo R.V.;
RT "The murine decorin. Complete cDNA cloning, genomic organization,
RT chromosomal assignment, and expression during organogenesis and
RT tissue differentiation.";
EL J. Biol. Chem. 269:28270-28281(1994).
CC -!- FUNCTION: May affect the rate of fibrils formation.
CC -!- SUBUNIT: Binds to type I and type II collagen, to fibronectin and
CC TGF-beta. Forms a ternary complex with MFAP2 and ELN (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Extracellular matrix.
CC -!- PTM: The attached glycosaminoglycan chain can be either
CC chondroitin sulfate or dermatan sulfate depending upon the tissue
CC of origin (By similarity).
CC -!- SIMILARITY: BELONGS TO THE SMALL LEUCINE-RICH PROTEOGLYCAN
CC (SLRPS) FAMILY. CLASS I SUBFAMILY.
CC -!- SIMILARITY: CONTAINS 12 LEUCINE-RICH REPEATS (LRR).

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X53929; CAA37876.1; -
DR PIR; S20812; S20812.
DR HSP; P23945; IXUN.
DR MGD; MGI:94872; Dcn.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR00372; LRR_Nterm.
DR InterPro; IPR003592; LRR_out.
DR InterPro; IPR003591; LRR_type.
DR Pfam; PF00560; LRR; 9.
DR SMART; PF01462; LRRNT; 1.
DR SMART; SM00370; LRR; 2.
DR SMART; SM00013; LRRNT; 1.
DR SMART; SM00369; LRR_type; 2.
KW Glycoprotein; Extracellular matrix; Proteoglycan; Repeat;
KW Leucine-rich repeat; Signal.
FT SIGNAL 1 16
FT PROPEP 17 30
FT CHAIN 31 354
FT DOMAIN 49 62
FT REPEAT 68 88
FT REPEAT 89 112
FT REPEAT 113 136
FT REPEAT 137 157
FT REPEAT 158 181
FT REPEAT 182 207
FT REPEAT 208 228
FT REPEAT 229 252
FT REPEAT 253 276
FT REPEAT 277 299
FT REPEAT 300 329
FT REPEAT 330 354
FT DISULFID 49 62
FT DISULFID 308 341
FT CARBOHYD 34 34
FT CARBOHYD 206 206
FT CARBOHYD 241 241
FT CARBOHYD 257 257
FT CARBOHYD 298 298
SQ SEQUENCE 354 AA; 39809 MW; F05B5CC08DCABF6F CRC64;

Query Match
Best Local Similarity 9.9%; Score 338; DB 1; Length 354;
Matches 91; Conservative 56; Mismatches 102; Indels 66; Gaps 9;

QY 9 FLIGTKTGLFQVAPLSVMAKSPVCRDAGFYCNDRFLSIPTGIPEDATHTYLNQNN 68
Db 27 FMLEDEASGIIPYDPDNPPLSMCPYRCQCHLVYVQCSDLGLDKVDFPDPDTLLDQNN 86
QY 69 QLN-----NAGI-----PSDLKLLKVERIYLYHNSLDEFTNPKY 105
Db 87 KITETKSGAFNKLKDLHTLLVNNKISKISPEAFKPLKRLYLSKQLKELPERPT 146
QY 106 VKELHLQENNI---RTTYDSKIPYLEELDDNSVSAVSEIEGAFRDSNYLRLLFLS 162
Db 147 LQELRVHNEITLKRKSDFNGLNNVLVIE---LGGNPKNSGIENGAFQGLKSLYSIRIS 203
QY 163 RNHLSTIPWGLPTIEELRLDNRLSISPSQGTSTSLKRLVLDGNNLNNHGLGDKVFF 222
Db 204 DTNITAIPOGLPTSLTEVHLDGNKTKYDAPSLKG----- 238
QY 223 NLYNLTLSIVRNSLT---AAPVNLPGTNRKLYLDHNNKINRVPN--AFSYLRQLYRL 276
Db 239 -LJNLKSLGSLSPNSITVMENGSLANVP--HLEHLDDNNKLLRVPAGLAQHKYIQVYV-- 293

```

QY 277 DMSNNLSNLPQGIIF 291
: |||: | | |
Db 294 -LHNNNISAVGQND 307

RESULT 12
PGS2_BOVIN STANDARD; PRT; 360 AA.
AC P21793;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Decorin precursor (Bone proteoglycan II) (PG-S2).
GN DCM.

OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=68133946; PubMed=3435485;
RA Day A.A., McQuillan C.I., Termini J.D., Young M.R.;
RT "Molecular cloning and sequence analysis of the cDNA for small
proteoglycan II of bovine bone.";
RL Biochem. J. 248:801-805(1987).
RN [2]
RP SEQUENCE OF 31-54.
RX MEDLINE=69123388; PubMed=2914936;
RA Choi H.U., Johnson T.L., Pal S., Tang L.H., Rosenberg L.C.,
RA Neame P.J.;
RT "Characterization of the dermatan sulfate proteoglycans, DS-PG1 and
DS-PGII, from bovine articular cartilage and skin isolated by octyl-
sepharose chromatography.";
RL J. Biol. Chem. 264:2876-2884(1989).
RN [3]
RP INTERACTIONS WITH MFAP2 AND ELN.
RX PubMed=11723132;
RA Reinboth B., Hansen E., Cleary E.G., Gibson M.A.;
RT "Molecular interactions of biglycan and decorin with elastic fiber
components: biglycan forms a ternary complex with tropoelastin and
microfibril-associated glycoprotein 1.";
RL J. Biol. Chem. 277:3950-3957(2002).
CC -!- FUNCTION: May affect the rate of fibrils formation.
CC -!- SUBUNIT: Binds to type I and type II collagen, to fibronectin and
TGF-beta. Forms a ternary complex with MFAP2 and ELN.
CC -!- SUBCELLULAR LOCATION: Extracellular matrix.
CC -!- PTM: The attached glycosaminoglycan chain can be either
chondroitin sulfate or dermatan sulfate depending upon the tissue
of origin.
CC -!- SIMILARITY: BELONGS TO THE SMALL LEUCINE-RICH PROTEOGLYCANS
(SLRPs) FAMILY. CLASS I SUBFAMILY.
CC -!- SIMILARITY: CONTAINS 12 LEUCINE-RICH REPEATS (LRR).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/))
or send an email to license@isb-sib.ch.
CC -----
CC EMBL; Y00712; CAA68702.1; -.
DR PIR; S06280; S06280.
DR PIR; B31430; B31430.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR000372; LRR Nterm.
DR InterPro; IPR003592; LRR Out.
DR InterPro; IPR003591; LRR Typ.
DR Pfam; PF00560; LRR; 8.
DR Pfam; PF01462; LRRNT; 1.
DR SMART; SM00370; LRR; 2.

DR SMART; SM00013; LRRNT; 1.
DR SMART; SM00369; LRR_TYP; 2.
KW Glycoprotein; Extracellular matrix; Proteoglycan; Repeat;
KW Leucine-rich repeat; Signal.
FT SIGNAL 1 16
FT PROPEP 17 30
FT CHAIN 31 360
FT DOMAIN 55 68
FT REPEAT 74 94
FT REPEAT 119 142
FT REPEAT 143 163
FT REPEAT 164 187
FT REPEAT 188 213
FT REPEAT 214 234
FT REPEAT 235 258
FT REPEAT 259 282
FT REPEAT 283 305
FT REPEAT 306 335
FT REPEAT 336 360
FT DISULFID 55 68
FT DISULFID 314 347
FT CARBOHYD 34 34
FT CARBOHYD 212 212
FT CARBOHYD 263 263
FT CARBOHYD 304 304
FT CARBOHYD 360 360
SQ SEQUENCE 360 AA; 39837 MW; 71E84DA2D87552C0 CRC64;

Query Match 9.9%; Score 336; DB 1; Length 360;
Best Local Similarity 28.8%; Pred. No. 1.9e-15;
Matches 92; Conservative 56; Mismatches 114; Indels 58; Gaps 8;
QY 18 FLOVAPLVSMAKSCVRCDCAGFYCNDRFTSTPTGIPEDATLYLQNNQINNA--GI 75
Db 42 FPEVPIEPGVCFCRCQCHLRVQCSDLGLEKVPKDLPPDTALLDLQNNKITEKGD 101
QY 76 PSDLKN-----LKVRIYLYHNSLDEFTNLKPKYKVELHLOEN 114
Db 102 FKNLKNLHTLILNNKISKISPGAFAPLVKLERLYLSKNQKLPKMKPTLQELRVHEN 161
QY 115 NRTIYDLSKIPYLELHDDNSYSAVSEEGAFEDSNLYRLLELSRNHLSTFWGILP 174
Db 162 EITVKRKSFTNGVOMIVVELGTNPGLASSIENGAFQGMKLSIRADNTITTIPOGLP 221
QY 175 RTBELRLDDNRISTSPSLOGTSLKRLVLDGNLANNHGLGDKVFNFNLNLTSLVR 234
Db 222 PSUTELHLDGNKITKVDAAELKGLNLAKLGLSFN-----SISAVD 262
QY 235 NSLTAAPVNLPGTNLRKLYLODNHINEVPEN--FSYLRQLYRLDMSNNLSNL----- 286
Db 263 NSLIA--NTP--HLREHLNANNKIAKPGGVADHKYIQVY---LHNNNISAGSNDFC 314
QY 287 PQGIFDDDLNITQLIRNPP 306
Db 315 PPGYNKKASYSVSLFSP 334

RESULT 13
PGS2_SHEEP STANDARD; PRT; 360 AA.
AC Q9TTE2;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Decorin precursor (Bone proteoglycan II) (PG-S2) (PG40).
GN DCM.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]

RP SEQUENCE FROM N.A.
RC TISSUE-Myometrium;
RX MEDLINE=20113292; PubMed=10644528;
RA Wu W.X., Zhang Q., Uno N., Derks J.B., Nathanielsz P.W.;
RT "Characterization of decorin mRNA in pregnant intrauterine tissues of
the ewe and regulation by steroids.";
RL Am. J. Physiol. 278:C199-C206(2000).
CC -!- FUNCTION: May affect the rate of fibrils formation (By
similarity).
CC -!- SUBUNIT: Binds to type I and type II collagen, to fibronectin and
TGF-beta. Forms a ternary complex with MFAP2 and ELN (By
similarity).
CC -!- SUBCELLULAR LOCATION: Extracellular matrix (By similarity).
CC -!- PTM: The attached glycosaminoglycan chain can be either
chondroitin sulfate or dermatan sulfate depending upon the tissue
of origin (By similarity).
CC -!- SIMILARITY: BELONGS TO THE SMALL LEUCINE-RICH PROTEOGLYCANS
(SLRPS) FAMILY. CLASS I SUBFAMILY.
CC -!- SIMILARITY: CONTAINS 12 LEUCINE-RICH REPEATS (LRR).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL: AF125041; AAF00585.1;
CC HSP: P09661; IAGN.
CC InterPro: IPR001611; LRR.
CC InterPro: IPR000372; LRR_Nterm.
CC InterPro: IPR003592; LRR_Out.
CC InterPro: IPR003591; LRR_Typ.
CC Pfam: PF00560; LRR; 9.
CC Pfam: PF01462; LRRNT; 1.
CC SMART: SM00370; LRR; 2.
CC SMART: SM00013; LRRNT; 1.
CC SMART: SM00369; LRR_Typ; 2.
CC Glycoprotein: Extracellular matrix; Proteoglycan; Repeat;
KW Leucine-rich repeat; Signal.
FT SIGNAL 1 16 POTENTIAL.
FT PROPEP 17 30 BY SIMILARITY.
FT CHAIN 31 360 DECORIN.
FT DOMAIN 55 68 CYS-RICH.
FT REPEAT 74 94 LRR-S 1.
FT REPEAT 95 118 LRR-T 1.
FT REPEAT 119 142 LRR-T 2.
FT REPEAT 143 163 LRR-S 2.
FT REPEAT 164 187 LRR-T 3.
FT REPEAT 188 213 LRR-T 4.
FT REPEAT 214 234 LRR-S 3.
FT REPEAT 235 258 LRR-T 5.
FT REPEAT 259 282 LRR-T 6.
FT REPEAT 283 305 LRR-S 4.
FT REPEAT 306 335 LRR-T 7.
FT REPEAT 336 360 LRR-T 8.
FT BY SIMILARITY.
FT DISULFID 55 68
FT DISULFID 314 347 O-LINKED (GLYCOSAMINOGLYCAN) (BY
FT CARBOHYD 34 34 SIMILARITY).
FT CARBOHYD 212 212 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 263 263 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 304 304 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 360 AA; 39972 MW; 0095DDDFAB86624 CRC64;
Query Match 9.9%; Score 336; DB 1; Length 360;
Best Local Similarity 29.1%; Pred. No. 1.9e-15;
Matches 93; Conservative 55; Mismatches 114; Indels 58; Gaps 8;
QY 18 FLOWAPISWAKSCSVRCDAHGYICNDFTSTIGTIPEDATLYLQNNQINNA--GI 75
DB 42 FHEVPELEPMGVCFCFCQCHLRVQCSDIGLEKVPKDLPPDTALLDQNNKITEIKGD 101

QY 76 PSDLN-----LLKVERIYLHNSLDSEPTNLPKYVKELHLEN 114
DB 102 FKNLNLHLIILINNKISKISPGAFAPVLEKLYLSKNQLKELPEKMPKTLOELRVHEN 161
QY 115 NIRITYDLSKIPYLEIEHLDDNSAVSIEGAFRDSNYIRLILFSRNHLSTIPWGLP 174
DB 162 EITKVRKSVFNGQIMIVVELGTNPFLKSSGIENGAFQGMKLSYIRIADTNITTIPOGLP 221
QY 175 RTIEELDDNRISITSPSLOGLTSLKRLVLDGNLNNHGLGDKVFNFNLVNLTELSYR 234
DB 222 PSLEFELHGNKTKYDAASLGLNLAKLGLSFN-----SISAYD 262
QY 235 NSLPAAPVNLPGTNLKRKYLODNEHNEVPNA--FSYLRQLYRLDMNNLSML----- 286
DB 263 NGSLA---NTP--HLRELHNNLNNKLVKVPGLADHKYIQVY---LHNNLSAIGSNDPC 314
QY 287 POGIFDDLDNITOLIIRNP 306
DB 315 PPGYNTKKASYSGVSLFSP 334
RESULT 14
PGS2-CORJA STANDARD; PRT; 356 AA.
AC Q9DE68;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Decorin precursor (Bone proteoglycan II) (PG-S2).
GN DCN.
OS Coturnix coturnix japonica (Japanese quail).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Coturnix.
OX NCBI TaxID=93934;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Cornea, and Sclera;
RX MEDLINE=20556471; PubMed=11102759;
RA Corpe L.M., Dunlevy J.R., Hassell J.R., Conrad A.H., Conrad G.W.;
RT "Molecular cloning and relative tissue expression of decorin and
lumican in embryonic quail cornea.";
RL Matrix Biol. 19:699-704(2000).
CC -!- FUNCTION: May affect the rate of fibrils formation (By
similarity).
CC -!- SUBUNIT: Binds to type I and type II collagen, to fibronectin and
TGF-beta. Forms a ternary complex with MFAP2 and ELN (By
similarity).
CC -!- SUBCELLULAR LOCATION: Extracellular matrix (By similarity).
CC -!- PTM: The attached glycosaminoglycan chain can be either
chondroitin sulfate or dermatan sulfate depending upon the tissue
of origin (By similarity).
CC -!- SIMILARITY: BELONGS TO THE SMALL LEUCINE-RICH PROTEOGLYCANS
(SLRPS) FAMILY. CLASS I SUBFAMILY.
CC -!- SIMILARITY: CONTAINS 12 LEUCINE-RICH REPEATS (LRR).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL: AF125250; AAG48154.1;
DR InterPro: IPR001611; LRR.
DR InterPro: IPR000372; LRR_Nterm.
DR InterPro: IPR003592; LRR_Out.
DR InterPro: IPR003591; LRR_Typ.
DR Pfam: PF00560; LRR; 10.
DR Pfam: PF01462; LRRNT; 1.
DR SMART: SM00370; LRR; 3.

```

DR SMART; SM00013; LRRNT; 1.
DR SMART; SM00369; LRR_TYP; 7.
KW Glycoprotein; Extracellular matrix; Proteoglycan; Repeat;
KW Leucine-rich repeat; Signal.
FT SIGNAL 1 15 POTENTIAL.
FT PROPEP 16 29 BY SIMILARITY.
FT CHAIN 30 356 DECORIN.
FT DOMAIN 51 64 CYS-RICH.
FT REPEAT 70 90 LRR-S 1.
FT REPEAT 91 114 LRR-T 1.
FT REPEAT 115 138 LRR-T 2.
FT REPEAT 139 159 LRR-S 2.
FT REPEAT 160 183 LRR-T 3.
FT REPEAT 184 209 LRR-T 4.
FT REPEAT 210 230 LRR-S 3.
FT REPEAT 231 254 LRR-T 5.
FT REPEAT 255 278 LRR-T 6.
FT REPEAT 279 301 LRR-S 4.
FT REPEAT 302 331 LRR-T 7.
FT REPEAT 332 356 LRR-T 8.
FT DISULFID 51 64 BY SIMILARITY.
FT DISULFID 310 343 BY SIMILARITY.
FT CARBOHYD 33 33 O-LINKED (GLYCOSAMINOGLYCAN) (BY
FT CARBOHYD 208 208 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 259 259 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 356 AA; 39571 MW; BE958306AED7DB26 CRC64;

Query Match 9.8%; Score 334; DB 1; Length 356;
Best Local Similarity 31.4%; Pred. No. 2.6e-15;
Matches 91; Conservative 51; Mismatches 88; Indels 60; Gaps 10;

QY 31 CPVCRDAGFYCNDRFLTSIPTGIPEDATLYLONNOINNA--GIPSDLKN----- 81
DB 51 CPFRQCHLRVWCSDGLGERVPKDLDPDTLLDQNNKITEIRGDFKNLKNLHALLIV 110
QY 82 -----LLKVERILYHNSLDEFFNLPKYVKEHLQENNI---RTIYDSL 124
DB 111 NKKISKISPOAFAPLKKLERLYLKNLKLPELNPMPKIQETRAHENEISKRAVENGL 170
QY 125 SKIPLYELHLDNSVANSIEGAFRDSNYLRLLFLSRNHLSTIPWGLPRTIEELRLDD 184
DB 171 NGVIVLE---LGNLPKSSGIEGAFQGMKLSYIRIATNTISPKGLPSFELHLDG 227
QY 185 NRISTSPSLOGITSLKRLVLDGNLNLHGLGDKVFFNLVNLTSLSVNRNSLTAAPVNL 244
DB 228 NKISKIDAEGSLGTNLAKLGLS-----FNSIS-----SVENGSLNVP--- 266
QY 245 PGTNLKLYLODNIHNVPP--NAPSYLRLQLYRLDMNSNNLNLPGQIFD 292
DB 267 ---HLRELHLNNELRVPSGLGEHKYIQVY---LHNNKIASI---GIND 308

RESULT 15
PGS2.CHICK STANDARD; PRT; 357 AA.
AC F28675;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Decorin precursor (Bone proteoglycan II) (PG-S2).
OS Gallus gallus (Chicken)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OC NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
RC STRAIN=White leghorn; TISSUE=Cornea;
RX MEDLINE=92296755; PubMed=1605630;
RA Li W., Vergnes J.-P., Cornuet P.K., Hassell J.R.;
RT "cDNA clone to chick corneal chondroitin/dermatan sulfate
RT proteoglycan reveals identity to decorin.";

```

```

RL Arch. Biochem. Biophys. 296:190-197(1992).
CC -I- FUNCTION: May affect the rate of fibrils formation (By
CC similarity).
CC -I- SUBUNIT: Binds to type I and type II collagen, to fibronectin and
CC TGF-beta. Forms a ternary complex with MEAP2 and ELN (By
CC similarity).
CC -I- SUBCELLULAR LOCATION: Extracellular matrix (By similarity).
CC -I- PTM: The attached glycosaminoglycan chain can be either
CC chondroitin sulfate or dermatan sulfate depending upon the tissue
CC of origin (By similarity).
CC -I- SIMILARITY: BELONGS TO THE SMALL LEUCINE-RICH PROTEOGLYCAN
CC (SLRPS) FAMILY. CLASS I SUBFAMILY.
CC -I- SIMILARITY: CONTAINS 12 LEUCINE-RICH REPEATS (LRR).
CC this SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X63797; CA443318.1; -
CC PIR; S22197; S22197.
CC PIR; S24317; S24317.
CC InterPro: IPR001611; LRR.
CC InterPro: IPR003372; LRR_Nterm.
CC InterPro: IPR003592; LRR_Out.
CC InterPro: IPR003591; LRR_Typ.
CC Pfam; PF00560; LRR; 8.
CC Pfam; PF01462; LRRNT; 1.
CC SMART; SM00370; LRR; 2.
CC SMART; SM00013; LRRNT; 1.
CC SMART; SM00369; LRR_Typ; 2.
KW Glycoprotein; Extracellular matrix; Proteoglycan; Repeat;
KW Leucine-rich repeat; Signal.
FT SIGNAL 1 16
FT PROPEP 17 30
FT CHAIN 31 357 DECORIN.
FT DOMAIN 52 65 CYS-RICH.
FT REPEAT 71 91 LRR-S 1.
FT REPEAT 92 115 LRR-T 1.
FT REPEAT 116 139 LRR-T 2.
FT REPEAT 140 160 LRR-S 2.
FT REPEAT 161 184 LRR-T 3.
FT REPEAT 185 210 LRR-T 4.
FT REPEAT 211 231 LRR-S 3.
FT REPEAT 232 255 LRR-T 5.
FT REPEAT 256 279 LRR-T 6.
FT REPEAT 280 302 LRR-S 4.
FT REPEAT 303 332 LRR-T 7.
FT REPEAT 333 357 LRR-T 8.
FT DISULFID 52 65 BY SIMILARITY.
FT DISULFID 311 344 BY SIMILARITY.
FT CARBOHYD 34 34 O-LINKED (GLYCOSAMINOGLYCAN) (BY
FT CARBOHYD 209 209 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 260 260 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 357 AA; 39687 MW; 31B104C7C3D711D CRC64;

Query Match 9.8%; Score 334; DB 1; Length 357;
Best Local Similarity 31.4%; Pred. No. 2.6e-15;
Matches 91; Conservative 51; Mismatches 88; Indels 60; Gaps 10;

QY 31 CPVCRDAGFYCNDRFLTSIPTGIPEDATLYLONNOIN--NAGIPSDLKN----- 81
DB 52 CPFRQCHLRVWCSDGLGERVPKDLDPDTLLDQNNKITEIRGDFKNLKNLHALLIV 111
QY 82 -----LLKVERILYHNSLDEFFNLPKYVKEHLQENNI---RTIYDSL 124
DB 112 NKKISKISPAAPLKKLERLYLKNLKLPELNPMPKSIQETRAHENEISKRAVENGL 171
QY 125 SKIPLYELHLDNSVANSIEGAFRDSNYLRLLFLSRNHLSTIPWGLPRTIEELRLDD 184

```

Db 172 NQVIVLE---LGTNPLKSGIENGAFQGMKRLSVIRIADTNITSIPKGLPPSLTELHLDG 228
QY 185 NRISTISSPSLOGLTSLKRLVLDGNNLNHGLGDKVFFNLVNLTELSLYRNSLTAAPVNL 244
Db 229 NKISKIDAEGLSGLTNLAKIGLS-----FNSIS----SVENGSLNNVP--- 267
QY 245 PGTNLRLKLYLODNHINRVP--NAPSFLRQLYRLDMNNLNLPOGIFD 292
Db 268 ---HLRELHANNELVVPGLGEHKYIQVY---LHNNKIASI--GIND 309

Search completed: April 23, 2003, 11:05:37
Job time : 18 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 11:03:31 ; Search time 37 Seconds
(without alignments)
3614.178 Million cell updates/sec

Title: US-10-004-551-6

Perfect score: 3400

Sequenced: 1 MISANWSILGKIGKIFLQ.....SSNRSYRDSGIPDSHSHS 649

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_21.*

1: sp-archaea.*

2: sp-bacteria.*

3: sp-fungi.*

4: sp-human.*

5: sp-invertebrate.*

6: sp-mammal.*

7: sp-mhc.*

8: sp-organelle.*

9: sp-phage.*

10: sp-plant.*

11: sp-rodent.*

12: sp-virus.*

13: sp-vertebrate.*

14: sp-unclassified.*

15: sp-rviro.*

16: sp-bacteriap.*

17: sp-archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3396	99.9	649	Q96KB1	Q96KB1 homo sapien
2	3395	99.9	662	Q9P259	Q9P259 homo sapien
3	3269	96.1	623	Q96K39	Q96K39 homo sapien
4	1944	57.2	372	Q96K42	Q96K42 homo sapien
5	1922	56.5	674	Q8WVA2	Q8WVA2 homo sapien
6	717.5	21.1	246	Q43408	Q43408 homo sapien
7	397	11.7	708	Q9H3W5	Q9H3W5 homo sapien
8	394.5	11.6	707	P97860	P97860 mus musculus
9	385	11.3	640	Q9HCJ2	Q9HCJ2 homo sapien
10	384	11.3	649	Q96A85	Q96A85 homo sapien
11	384	11.3	653	Q9HBM1	Q9HBM1 homo sapien
12	382.5	11.2	1523	Q9WVB4	Q9WVB4 mus musculus
13	361.5	11.2	1523	Q8E280	Q8E280 rattus norv
14	378	11.1	648	Q8VI35	Q8VI35 mus musculus
15	376.5	11.1	707	Q9E5Y6	Q9E5Y6 rattus norv
16	374	11.0	1523	Q75094	Q75094 homo sapien

17	370.5	10.9	716	11	Q61809
18	364.5	10.7	581	4	Q8TF66
19	360.5	10.6	718	13	Q73675
20	359.5	10.6	1480	5	Q9V7F8
21	359.5	10.6	1504	5	Q9V7F9
22	358.5	10.5	578	11	Q8E5M3
23	356.5	10.5	1504	5	Q9XIV4
24	353.5	10.4	730	4	Q9P231
25	351.5	10.3	1515	13	Q9DEB37
26	348.5	10.2	1534	4	Q75093
27	347.5	10.2	1531	11	Q88279
28	345.5	10.2	1521	4	Q95710
29	345.5	10.2	1529	4	Q94813
30	345.5	10.2	1530	13	Q90WZ3
31	344.5	10.1	1521	11	Q9R1B9
32	343.5	10.1	1512	13	Q9DEB36
33	343.5	10.1	1525	4	Q9Y5Q7
34	342.5	10.1	1531	11	Q9WVB5
35	341.5	10.0	1530	11	Q9WUG5
36	334.5	9.8	359	13	Q9DE03
37	333	9.8	562	11	Q8R113
38	333	9.8	570	11	Q9DBB9
39	332.5	9.8	673	11	Q9CZT5
40	332.5	9.8	798	4	Q8WVZ2
41	331.5	9.8	673	11	Q8R2G5
42	330	9.7	1094	4	Q9BYB8
43	329.5	9.7	796	11	Q9WVC1
44	327	9.6	581	6	Q9BGP6
45	326	9.6	581	6	Q95KI8

ALIGNMENTS

RESULT 1

Q96KB1 ID Q96KB1 PRELIMINARY; PRT; 649 AA.

AC Q96KB1; TISSUE=EMBRYO, AND MAINLY HEAD;

DT 01-DEC-2001 (TREMREL. 19, Created)

DT 01-DEC-2001 (TREMREL. 19, Last sequence update)

DT 01-MAR-2002 (TREMREL. 20, Last annotation update)

DE CDNA FLJ14391 fis, clone HEMBA1003077, weakly similar to SLIT protein precursor.

DE Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RM [1]

RP SEQUENCE FROM N.A.

RC TISSUE=EMBRYO, AND MAINLY HEAD;

RA Isoyagi T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y., RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H., RA Wagatsuma M., Hosoi T., Kaku Y., Kodaira H., Kondo H., Sugawara M., RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S., RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K., RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y., RA Ninomiya K., Iwayanagi T.;

RT "NEDO human cDNA sequencing project.";

RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.

EMBL; AK027297; BAB55023.1; -

DR InterPro; IPR003961; FN.III.

DR InterPro; IPR001611; LRR.

DR InterPro; IPR000483; LRR_Cterm.

DR InterPro; IPR000372; LRR_Nterm.

DR InterPro; IPR001211; PhospholipaseA2.

DR Pfam; PF00041; fn3; 1.

DR Pfam; PF00560; LRR; 10.

DR Pfam; PF01463; LRRCT; 1.

DR Pfam; PF01462; LRRNT; 1.

DR	SMART; SM00060; FN3; 1.
DR	SMART; SM00013; LRRNT; 1.
DR	PROSITE; PS00119; PA2_ASP; UNKNOWN_1.
SQ	SEQUENCE 649 AA; 72994 MW; 820B527E53F62F08 CRC64;

Query Match 99.9%; Score 3396; DB 4; Length 649;
Best Local Similarity 99.8%; Pred. No. 3.1e-242;
Matches 648; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MISAAWIFLIGTKIGLFLQVAPLSVMAKSCPSVCRDAGFYCNDRLFTSIPTGIPEDA 60
DB 1 MISAAWIFLIGTKIGLFLQVAPLSVMAKSCPSVCRDAGFYCNDRLFTSIPTGIPEDA 60

QY 61 TTYLQNNQINNAGISDLSKLLKVERIYLYHNSLDEFTNLPKYVKELHQLQNNIRIT 120
DB 61 TTYLQNNQINNAGISDLSKLLKVERIYLYHNSLDEFTNLPKYVKELHQLQNNIRIT 120

QY 121 YDSLSKIPYLEELHDDNSVSAVSEEGAFRDSNYLRLLFLSRNHLSTIPWGLPRTIEEL 180
DB 121 YDSLSKIPYLEELHDDNSVSAVSEEGAFRDSNYLRLLFLSRNHLSTIPWGLPRTIEEL 180

QY 181 RLDNRISTISPSLQGLTSLKRLVLDGNLLNHHGLGDKVFFNLVNLTELSTVNSLITAA 240
DB 181 RLDNRISTISPSLQGLTSLKRLVLDGNLLNHHGLGDKVFFNLVNLTELSTVNSLITAA 240

QY 241 PVNLPGLNLRKLYLDQNHINRVPNAPSYLRQLYRDLMSNNLSNLPQGFDDLDNITQL 300
DB 241 PVNLPGLNLRKLYLDQNHINRVPNAPSYLRQLYRDLMSNNLSNLPQGFDDLDNITQL 300

QY 301 ILRNNPWYCGCKMKWVDMQSLPVKNVRLGMCQAPKVRGMAIKDLNAELFDCKDSGI 360
DB 301 ILRNNPWYCGCKMKWVDMQSLPVKNVRLGMCQAPKVRGMAIKDLNAELFDCKDSGI 360

QY 361 VSTIQITTAIPNTVPAQGWPAVTKQPKDKNPKLTKDQQTGSPSRKTIITTVKSVTS 420
DB 361 VSTIQITTAIPNTVPAQGWPAVTKQPKDKNPKLTKDQQTGSPSRKTIITTVKSVTS 420

QY 421 DTIHSWKALPMTALRLSWLKLGHSPAGSITETIVTGERSEYLVTALEPDSYKVCW 480
DB 421 DTIHSWKALPMTALRLSWLKLGHSPAGSITETIVTGERSEYLVTALEPDSYKVCW 480

QY 481 PMETSNIYLFDETPVCIEETAPLRMNPNTTLNREQEKPYKNPNLPLAAIIGGAVLV 540
DB 481 PMETSNIYLFDETPVCIEETAPLRMNPNTTLNREQEKPYKNPNLPLAAIIGGAVLV 540

QY 541 TIALALVWVVRNGLSFRNCAYSKGRRKDDYAEAGTKKDNSILIRETSFQMLPTS 600
DB 541 TIALALVWVVRNGLSFRNCAYSKGRRKDDYAEAGTKKDNSILIRETSFQMLPTS 600

QY 601 NEPISKEEFVHTIFPPNGMNLKNNHSESSNRSYRDSGIPDSHSHS 649
DB 601 NEPISKEEFVHTIFPPNGMNLKNNHSESSNRSYRDSGIPDSHSHS 649

RESULT 2
Q9P259 PRELIMINARY; PRT; 662 AA.
AC Q9P259;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE KIAA1469 protein (Fragment).
GN KIAA1469.

OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20277482; PubMed=10819331;
RA Nagase T., Kikuno R., Ishikawa K., Hirose M., Ohara O.;
RT "Prediction of the coding sequences of unidentified human
RT genes.XVII.the complete sequences of 100 new cDNA clones from brain
RT which code for large proteins in vitro.";
RL DNA Res. 7:143-150(2000).
DR EMBL; AB040902; BAB5993.1; -.
DR InterPro; IPR003961; FN_III.

InterPro; IPR001611; LRR.
DR InterPro; IPR000483; LRR_Cterm.
DR InterPro; IPR000372; LRR_Nterm.
DR InterPro; IPR003592; LRR_Out.
DR InterPro; IPR003591; LRR_Typ.
DR InterPro; IPR001211; PhospholipaseA2.
DR Pfam; PF00041; fn3; 1
DR Pfam; PF00560; LRR; 10.
DR Pfam; PF01463; LRRCT; 1.
DR Pfam; PF01462; LRRNT; 1.
DR SMART; SM00060; FN3; 1.
DR SMART; SM00370; LRR; 1.
DR SMART; SM00082; LRRCT; 1.
DR SMART; SM00013; LRRNT; 1.
DR SMART; SM00369; LRR_Typ; 1.
DR PROSITE; PS00119; PA2_ASP; UNKNOWN_1.
FT NON_TER 1
SQ SEQUENCE 662 AA; 74459 MW; FF7CCCE689E4B429 CRC64;

Query Match 99.9%; Score 3395; DB 4; Length 662;
Best Local Similarity 99.8%; Pred. No. 3.8e-242;
Matches 648; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MISAAWIFLIGTKIGLFLQVAPLSVMAKSCPSVCRDAGFYCNDRLFTSIPTGIPEDA 60
DB 1 MISAAWIFLIGTKIGLFLQVAPLSVMAKSCPSVCRDAGFYCNDRLFTSIPTGIPEDA 73

QY 61 TTYLQNNQINNAGISDLSKLLKVERIYLYHNSLDEFTNLPKYVKELHQLQNNIRIT 120
DB 74 TTYLQNNQINNAGISDLSKLLKVERIYLYHNSLDEFTNLPKYVKELHQLQNNIRIT 133

QY 121 YDSLSKIPYLEELHDDNSVSAVSEEGAFRDSNYLRLLFLSRNHLSTIPWGLPRTIEEL 180
DB 134 YDSLSKIPYLEELHDDNSVSAVSEEGAFRDSNYLRLLFLSRNHLSTIPWGLPRTIEEL 193

QY 181 RLDNRISTISPSLQGLTSLKRLVLDGNLLNHHGLGDKVFFNLVNLTELSTVNSLITAA 240
DB 194 RLDNRISTISPSLQGLTSLKRLVLDGNLLNHHGLGDKVFFNLVNLTELSTVNSLITAA 253

QY 241 PVNLPGLNLRKLYLDQNHINRVPNAPSYLRQLYRDLMSNNLSNLPQGFDDLDNITQL 300
DB 254 PVNLPGLNLRKLYLDQNHINRVPNAPSYLRQLYRDLMSNNLSNLPQGFDDLDNITQL 313

QY 301 ILRNNPWYCGCKMKWVDMQSLPVKNVRLGMCQAPKVRGMAIKDLNAELFDCKDSGI 360
DB 314 ILRNNPWYCGCKMKWVDMQSLPVKNVRLGMCQAPKVRGMAIKDLNAELFDCKDSGI 373

QY 361 VSTIQITTAIPNTVPAQGWPAVTKQPKDKNPKLTKDQQTGSPSRKTIITTVKSVTS 420
DB 374 VSTIQITTAIPNTVPAQGWPAVTKQPKDKNPKLTKDQQTGSPSRKTIITTVKSVTS 433

QY 421 DTIHSWKALPMTALRLSWLKLGHSPAGSITETIVTGERSEYLVTALEPDSYKVCW 480
DB 434 DTIHSWKALPMTALRLSWLKLGHSPAGSITETIVTGERSEYLVTALEPDSYKVCW 493

QY 481 PMETSNIYLFDETPVCIEETAPLRMNPNTTLNREQEKPYKNPNLPLAAIIGGAVLV 540
DB 494 PMETSNIYLFDETPVCIEETAPLRMNPNTTLNREQEKPYKNPNLPLAAIIGGAVLV 553

QY 541 TIALALVWVVRNGLSFRNCAYSKGRRKDDYAEAGTKKDNSILIRETSFQMLPTS 600
DB 554 TIALALVWVVRNGLSFRNCAYSKGRRKDDYAEAGTKKDNSILIRETSFQMLPTS 613

QY 601 NEPISKEEFVHTIFPPNGMNLKNNHSESSNRSYRDSGIPDSHSHS 649
DB 614 NEPISKEEFVHTIFPPNGMNLKNNHSESSNRSYRDSGIPDSHSHS 662

RESULT 3
Q96K39 PRELIMINARY; PRT; 623 AA.
ID Q96K39;
AC Q96K39;
DT 01-DEC-2001 (TReMBLrel. 19, Created)

```

DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE CDNA FLJ14788 fis, clone NT2RP4000925, weakly similar to fibromodulin
DE precursor.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Isegai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Aotsuka S., Yoshikawa Y.,
RA Matsunawa H., Ishii S., Kawai Y., Saito K., Yamamoto J., Wakamatsu A.,
RA Nakamura Y., Nagahari K., Masuho Y., Sasaki N.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK027694; BAB55303.1; -
DR InterPro; IPR003961; FN.III.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR000483; LRR_Cterm.
DR InterPro; IPR000372; LRR_Nterm.
DR InterPro; IPR001211; PhospholipaseA2.
DR Pfam; PF00041; fn3; 1.
DR Pfam; PF00560; LRR; 10.
DR Pfam; PF01463; LRRCT; 1.
DR Pfam; PF01462; LRRNT; 1.
DR SMART; SM00060; FN3; 1.
DR SMART; SM00013; LRRNT; 1.
DR PROSITE; PS00119; FA2_ASP; UNKNOWN_1.
SQ SEQUENCE 623 AA; 70245 MW; F3BAG957CFID7A84 CRC64;

Query Match 96.1%; Score 3269; DB 4; Length 623;
Best Local Similarity 99.8%; Pred. No. 7e-233;
Matches 622; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 27 MAKSPSVRCDAAGFIYCNDRFLTSTPGICPEDATTLYLQNNQINNAGIPSLDKNLKVE 86
DB 1 MAKSPSVRCDAAGFIYCNDRFLTSTPGICPEDATTLYLQNNQINNAGIPSLDKNLKVE 60

QY 87 RIVYHNSLDFFNLPKYKELHLOENIRITVDSLSKIPYLEELHDDNSVSANSIE 146
DB 61 RIVYHNSLDFFNLPKYKELHLOENIRITVDSLSKIPYLEELHDDNSVSANSIE 120

QY 147 EGAFRDSNVLRLFLSRNHLSTIPWGLPTTIEELRLDNRISTISPSLQGLTSLKRLVL 206
DB 121 EGAFRDSNVLRLFLSRNHLSTIPWGLPTTIEELRLDNRISTISPSLQGLTSLKRLVL 180

QY 207 DGMNLNHEGLGDKVFFNLVNLTELSVRNSLTAAPVNLPGTNLRKLYLQDNHINRVPNA 266
DB 181 DGMNLNHEGLGDKVFFNLVNLTELSVRNSLTAAPVNLPGTNLRKLYLQDNHINRVPNA 240

QY 267 FSTLRQLYRLDMSNNLSNLPQGIFFDDLDNITQLILRNNPWYCGCKMKVROWLQSLPVK 326
DB 241 FSTLRQLYRLDMSNNLSNLPQGIFFDDLDNITQLILRNNPWYCGCKMKVROWLQSLPVK 300

QY 327 VNVYGLMCAQPKVRGMAIKDLNAELFDCKDSGIVSTIQTITAPNTVYPAQOWPAPVT 386
DB 301 VNVYGLMCAQPKVRGMAIKDLNAELFDCKDSGIVSTIQTITAPNTVYPAQOWPAPVT 360

QY 387 KQPDINKPKLTKDQQTGSPSRKTIITVKSVTSDTIHISWKLALPMTALRLSWLKLGH 446
DB 361 KQPDINKPKLTKDQQTGSPSRKTIITVKSVTSDTIHISWKLALPMTALRLSWLKLGH 420

QY 447 PAFGSGITETIVGSESELYTALPDPSPKYKCMVPMETSNLYLDFETFPVCIETETAPLRM 506
DB 421 PAFGSGITETIVGSESELYTALPDPSPKYKCMVPMETSNLYLDFETFPVCIETETAPLRM 480

QY 507 YNPFTTLNREQEKEPKNPNPLAAIIGGVALVTIALLVGVYVHRNGSLFSRNCAYS 566
DB 481 YNPFTTLNREQEKEPKNPNPLAAIIGGVALVTIALLVGVYVHRNGSLFSRNCAYS 540

QY 567 KGRRKDDYAEAGTKKNSILEIRETSFQMLPISNEPISKEEFVIHTIFFPPNGMNLKNN 626
DB 541 KGRRKDDYAEAGTKKNSILEIRETSFQMLPISNEPISKEEFVIHTIFFPPNGMNLKNN 600

QY 627 HSESSNRSYRSDGIPDSHSHS 649
DB 601 HSESSNRSYRSDGIPDSHSHS 623

RESULT 4
Q96K42 PRELIMINARY; PRT; 372 AA.
AC Q96K42;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE CDNA FLJ14764 fis, clone NT2RP3003716, weakly similar to SLIT protein
DE precursor.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Isegai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
RA Wadatsuma M., Hosoiiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y.,
RA Ninomiya K., Iwayanagi T.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK027670; BAB55282.1; -
DR InterPro; IPR003961; FN.III.
DR InterPro; IPR000483; LRR_Cterm.
DR Pfam; PF00041; fn3; 1.
DR Pfam; PF01463; LRRCT; 1.
DR SMART; SM00060; FN3; 1.
SQ SEQUENCE 372 AA; 41614 MW; A45B2E0E729AB69F CRC64;

Query Match 57.2%; Score 1944; DB 4; Length 372;
Best Local Similarity 99.2%; Pred. No. 2.2e-135;
Matches 369; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 278 MSNNLSNLPQGIFFDDLDNITQLILRNNPWYCGCKMKVROWLQSLPVKNVRLMCAQ 337
DB 1 MSNNLSNLPQGIFFDDLDNITQLILRNNPWYCGCKMKVROWLQSLPVKNVRLMCAQ 60

QY 338 EKVRGMAIKDLNAELFDCKDSGIVSTIQTITAPNTVYPAQOWPAPVTKPDINKPKLT 397
DB 61 EKVRGMAIKDLNAELFDCKDSGIVSTIQTITAPNTVYPAQOWPAPVTKPDINKPKLT 120

QY 398 KQQTGSPSRKTIITVKSVTSDTIHISWKLALPMTALRLSWLKLGHSPAGFSITETIV 457
DB 121 KQQTGSPSRKTIITVKSVTSDTIHISWKLALPMTALRLSWLKLGHSPAGFSITETIV 180

QY 458 TGERSELYTALPDPSPKYKCMVPMETSNLYLDFETFPVCIETETAPLRMNTTLNRQ 517
DB 181 TGERSELYTALPDPSPKYKCMVPMETSNLYLDFETFPVCIETETAPLRMNTTLNRQ 240

QY 518 KEPEYKPNPLAAIIGGVALVTIALLVGVYVHRNGSLFSRNCAYSKGRRKDDYAE 577
DB 241 KEPEYKPNPLAAIIGGVALVTIALLVGVYVHRNGSLFSRNCAYSKGRRKDDYAE 300

QY 578 AGTKKNSILEIRETSFQMLPISNEPISKEEFVIHTIFFPPNGMNLKNNHSESSNSR 637
DB 301 AGTKKNSILEIRETSFQMLPISNEPISKEEFVIHTIFFPPNGMNLKNNHSESSNSR 360

QY 638 DSGIPDSHSHS 649
DB 361 DSGIPDSHSHS 372

RESULT 5

```

```
Q8WVA2
ID Q8WVA2 PRELIMINARY; PRT; 674 AA.
AC Q8WVA2;
DT 01-MAR-2002 (TEMBLrel. 20, Created)
DT 01-MAR-2002 (TEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TEMBLrel. 21, Last annotation update)
DE Hypothetical 74.1 kDa protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-BRAIN;
RA Strausberg R.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DDJ databases.
DR EMBL; BC018370; AAH18370.1; -.
DR InterPro; IPR003961; FN.III.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR000483; LRR_Cterm.
DR InterPro; IPR000372; LRR_Nterm.
DR InterPro; IPR003592; LRR_out.
DR InterPro; IPR003591; LRR_Typ.
DR InterPro; IPR001211; PhospholipaseA2.
DR Pfam; PF00041; fn3; 1.
DR Pfam; PF00560; LRR; 7.
DR Pfam; PF01463; LRRCT; 1.
DR Pfam; PF01482; LRRNT; 1.
DR PRINTS; PR00019; LEURICHRPT.
DR SMART; SM00370; LRR; 3.
DR SMART; SM00082; LRRCT; 1.
DR SMART; SM00013; LRRNT; 1.
DR SMART; SM00369; LRR_TVP; 8.
DR PROSITE; PS00119; PA2_ASP; UNKNOWN_1.
KW Hypothetical protein.
SQ SEQUENCE 674 AA; 74087 MW; 38AB53F7243166CC CRC64;

Query Match 56.5%; Score 1922.5; DB 4; Length 674;
Best Local Similarity 58.7%; Pred. No. 1.9e-133;
Matches 382; Conservative 92; Mismatches 160; Indels 17; Gaps 9;

QY 6 NSIFLIGTKIGLQVAPISVMAKSCPSVCRDAGFYCNDRFLSITGIPEDATLYL 65
DB 34 W-LFICGLIAFTEV-----IDSTCPSPVCRDNGFYCNDRGLSIPADIPDATTLYL 88
QY 66 QNNQINAGIPSDIKNLLKVERIYLYHNSLDEFTNLPKYKELHLOENNIITTYDSLS 125
DB 89 QNNQINAGIPQDLKTKVNVQVYLYENDLDEFINLPRLSRELHLQDNVVTIARSLA 148
QY 126 KIPYLEELHDDNSVSAVSEIACAPRDSNYRLFLSRNHLSTIPWGLPRTIEBLRDDN 185
DB 149 RIPLLEKILHDDNSVSTVSEIDAFADSKQLKLLSRNHLSSIPSGLPHTLEELRDDN 208
QY 186 RISTISSPQLGTSIKRLVLDGNLLNHHGLDKVPFNVLNLTLSLVNSLTAAPVNL 245
DB 209 RISTIPHLAFKGLNSRLVLDGNLLANQRIADTFSLRNLTELSLVNSLAAPVNL 268
QY 246 GTNRLKYLQDNHINRVPNPFASYLRQLYRLDMSNNLSNLPQIGFDLDNITQLIRNN 305
DB 269 SAHLQKLYQDNATSHIPYNTLAKARELERLDLSNNNLATLPRGLFDLGNLAQLLRNN 328
QY 306 PNYCCCKMKVROKLOSLPKVNVGLMCOAPEKVRGMAIKDLNLELFDCKDS---GIV 361
DB 329 PWFCCGNCMLWLDVWVKAQAVNVNGLMCOGPEKVRGMAIKDITSEDECFTQPGGVA 388
QY 362 SYTIQTITAPN-TVYPAGQGWAPVTKQPKDKNPKLFDQQTGSPSRKTTITIVKSVTS 420
DB 389 NAAAKTASNASATTPQGSILFTLAKRPLRLPDSNIDYPMATGDAKTLAIHKALTA 448
QY 421 DTHISWKLALPMTALRLSKLGHSPAFGSGITITVGRSEYLVTALEPDSPKYKVM 480
DB 449 DSIRITWKATLPASSFRLSLRGLHSPAVGSGITITVGRSEYLVTALEPDSPKYKVM 508
```

```
QY 481 PMTSMNLYLDETPVCIETETAPLRMNPNTTILNREQKEPKYKPNPLAALIIGAVLV 540
DB 509 TMETSNAYADETPVCAKAETAD--SYGPTTILNREQNAGPM--ASLPLAGIIGAVLV 564
QY 541 TIAL-LALVGVWVHRNGSLFRMCAYSKGRKKDDYADAGTKKDNLSLEIRETSFQMLPI 599
DB 565 FLFLVGLAICWYVHAGELLTRERAYNRGRKKDDYMGSGTKKDNLSLEIRFGQLMPLI 624
QY 600 SNPEISKEFEVHIIFPPNGMNLKNNHSES--SSNRSTROSGIPDSDSHS 649
DB 625 -NPRAKEEYVHVHTIFPSNGSLCKATHTIGYGTTRGYRGGIPDIDYSY 674

RESULT 6
ID Q43408 PRELIMINARY; PRT; 246 AA.
AC Q43408;
DT 01-JUN-1998 (TremBLrel. 06, Created)
DT 01-JUN-1998 (TremBLrel. 06, Last sequence update)
DT 01-JUN-2001 (TremBLrel. 17, Last annotation update)
DE Hypothetical 26.6 kDa protein (fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-BRAIN;
RX MEDLINE=96207227; PubMed=8619474;
RA Andersson B., Wentland M.A., Ricafrente J.Y., Liu W., Gibbs R.A.;
RT "A 'double adaptor' method for improved shotgun library construction.";
RL Anal. Biochem. 236:107-113(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE-BRAIN;
RX MEDLINE=97264341; PubMed=9110174;
RA Yu W., Andersson B., Worley K.C., Muzny D.M., Ding Y., Liu W.,
RA Ricafrente J.Y., Wentland M.A., Lennon G., Gibbs R.A.;
RT "Large-scale concatenation cDNA sequencing.";
RL Genome Res. 7:353-358(1997).
DR EMBL; AF007139; AAC19151.1; -.
DR InterPro; IPR003961; FN.III.
DR Pfam; PF00041; fn3; 1.
DR KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 246 AA; 26646 MW; 9748DA32A74D461F CRC64;

Query Match 21.1%; Score 717.5; DB 4; Length 246;
Best Local Similarity 59.3%; Pred. No. 4.1e-45;
Matches 144; Conservative 36; Mismatches 56; Indels 7; Gaps 5;

QY 409 KTIITVTKSVTSDTHISWKLALPMTALRLSKLGHSPAFGSGITITVGRSEYLVTA 468
DB 9 KTLAIHKALTADTSIRITWKATLPASSFRLSLRGLHSPAVGSGITITVGRSEYLVTA 68
QY 469 LEPSDKYKVMFPMSTSNLYLDETPVCIETETAPLRMNPNTTILNREQKEPKYKPNLP 528
DB 69 LEPSKYYIICWVMTSMNAYADETPVCAKAETAD--SYGPTTILNREQNAGPM--ASLP 124
QY 529 LAATIGGVALVTIAL-LALVGVWVHRNGSLFRMCAYSKGRKKDDYADAGTKKDNLSL 587
DB 125 LAGIIGGVALVFLVGLAICWYVHAGELLTRERAYNRGRKKDDYMGSGTKKDNLSL 184
QY 588 EIRETSFQMLPISNPEISKEFEVHIIFPPNGMNLKNNHSES--SSNRSTROSGIPDSDSH 646
DB 185 EIRPGQLMPLI-NPYRAKEEYVHVHTIFPSNGSLCKATHTIGYGTTRGYRGGIPDIDY 243
QY 647 SHS 649
DB 244 SYT 246
```

```

RESULT 7
Q9H3W5 PRELIMINARY; PRT; 708 AA.
AC Q9H3W5; 043377;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical 79.4 kDa protein (Neuronal leucine-rich repeat protein-3).
GN DKF2P761K2424 OR NLRR-3 OR RGL18D07.1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RN [2]
RN [3]
RP SEQUENCE FROM N.A.
RA Hamano S., Inuzuka H., Morohashi A., Ohira M., Nakagawara A.;
RT "Human neuronal leucine-rich repeat protein-3(NLRR-3).";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 4-708 FROM N.A.
RC TISSUE=FIBROBLAST;
RA Murray J., Langston Y., Clarke C.;
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL442092; CAC09450.1; -
DR EMBL; AB060967; BAB47184.1; -
DR EMBL; AC004142; AAC02752.1; -
DR InterPro; IPR003961; FN_III.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003598; Ig_C2.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR000483; LRR_Cterm.
DR InterPro; IPR000372; LRR_Nterm.
DR InterPro; IPR003592; LRR_out.
DR InterPro; IPR003591; LRR_typ.
DR Pfam; PF00041; fn3; 1.
DR Pfam; PF00047; Ig; 1.
DR Pfam; PF00560; LRR; 8.
DR Pfam; PF01463; LRRCT; 1.
DR Pfam; PF01462; LRRNT; 1.
DR PRINTS; PR00019; LEURICHRPT.
DR SMART; SM00060; FN3; 1.
DR SMART; SM00409; IG; 1.
DR SMART; SM00408; IGG2; 1.
DR SMART; SM00370; LRR; 4.
DR SMART; SM00082; LRRCT; 1.
DR SMART; SM00013; LRRNT; 1.
DR SMART; SM00369; LRR_typ; 8.
KW Hypothetical protein; Immunoglobulin domain.
SQ SEQUENCE 708 AA; 79424 MW; 24710478D6B124D1 CRC64;

Query Match 11.7%; Score 397; DB 4; Length 708;
Best Local Similarity 23.2%; Pred. NO. 8.4e-21;
Matches 173; Conservative 114; Mismatches 240; Indels 218; Gaps 35;

QY 10 LIGTKIGLEQVAPISVWAKSCPSVREC-----DAGFYCNDRETFTIPTGI 56
DB 11 LGLAITTVQVADKKV---DCPRLCCTCEIRPWFPTFSIYMEASTVDCNDLGLLTFPARK 67
QY 57 PEDATTVYLNQNOI-----NNAGIPSDKLNLLKVERI-----YLYENS 94
DB 68 PANTQILLQTNNAIAKIEYSTDPEVNLTGDLISQNNLSVTVNINVKMPQLLSVLEENK 127
QY 95 LDEFP-----TNLPK-YVK-----ELHQENNTITFTYDLSKI 127
DB 128 LTELPEKCLSELNLOELYINHNLLSTISPGAFIGHNLLRLHLNSNRLQMINSKWFDAL 187

```

```

QY 128 PYLELHLDNDSYSAVSIEGAFRDSNVRLRLFLSRNHLSTIP-----WGLPRTTEELRLD 183
DB 188 PNLEILMIGENPI--TRIKDMFKPLINRLSLVLAGINTEIPDNALVGL-ENLESISFY 244
QY 184 DNRISTSSPSLQGLTSLKRLVLDGNLNNHGLGDKVFNFLNVLNLTSLVNSL-----237
DB 245 DNRLIKVPHVALQKVYNLAKFLDLNKNPINRIRGCD--FSNMLHLKELGI--NNMPELISI 300
QY 238 -TAPVNLPGTNLRKLYLODN-HINKVPPNAPSYLRQLYLDMSNNNLSNLPQGFIDLD 295
DB 301 DSLAVDNLDP--DLRKTEATNPNRLSYIHPNAPFLPKLESMLNSNLSALYHGTIESLP 358
QY 296 NITOLILRNPNWCGCKMKWVRDNLQSLPVKVVNR-----GLMCOAPEKVKVGMK----346
DB 359 NLKEISHTSNPIRCDCVIRWVN-----MKNTRNFMEDPSLFCVDPPEFQGVNRQVHF 412
QY 347 -----DINAEI-----FDCRSGIVSTIQ-----ITTA-----IPN 372
DB 413 RDMIEICILPIAPESPSPNLNVEAGSYVSFPCR-----ATAEPQPEIYWIITPSGQKLLPN 467
QY 373 T-----VYRAGQWPAPVTKQ---PDIKNPKLKD---QQTGSPSR 408
DB 468 TITDFYVHSEGLDINGVTPKEGLYTCIATNLVGADLASVMIKVDGSGFFQDNNGS---524
QY 409 KTIITTVKSVTSDTIHISWKLALPMTALRLSW---LKLHSPAFGSITETIVTGERSEYL 465
DB 525 --LNKIRDIQANSVLVSKASSKILKSSVKWTAFAVKTENSAAQASAR---IPSDVKVYN 579
QY 466 VTALPDPSPKVKVMPMEVSNLYLDFEIPVCIETETAPLRYNPTTLNRQEKPEYKNP 525
DB 580 LPHLPSTEXKICI-----DIPTIYQNRK--KCVNVTKGLHPDQKEYEKNK 625
QY 526 NLPLAAIIGGAVALVTIALIALVCMVVRNGLSFS--RNC-----AYSKGRRKKDYA---576
DB 626 TTTLMACLGGLGIIGVICLI-----SCLSPENMCDGHSYVRNLYQKTFALGE 675
QY 577 -----EAGTKKNSILEIRET 592
DB 676 LYPLINLWEAGKEKSTS-LKVKAT 699

RESULT 8
P97860 PRELIMINARY; PRT; 707 AA.
ID P97860
AC P97860;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Leucine-rich repeat protein precursor (Fragment).
GN LRRN3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RN [2]
RN [3]
RN [4]
RN [5]
RN [6]
RN [7]
RN [8]
RN [9]
RN [10]
RN [11]
RN [12]
RN [13]
RN [14]
RN [15]
RN [16]
RN [17]
RN [18]
RN [19]
RN [20]
RN [21]
RN [22]
RN [23]
RN [24]
RN [25]
RN [26]
RN [27]
RN [28]
RN [29]
RN [30]
RN [31]
RN [32]
RN [33]
RN [34]
RN [35]
RN [36]
RN [37]
RN [38]
RN [39]
RN [40]
RN [41]
RN [42]
RN [43]
RN [44]
RN [45]
RN [46]
RN [47]
RN [48]
RN [49]
RN [50]
RN [51]
RN [52]
RN [53]
RN [54]
RN [55]
RN [56]
RN [57]
RN [58]
RN [59]
RN [60]
RN [61]
RN [62]
RN [63]
RN [64]
RN [65]
RN [66]
RN [67]
RN [68]
RN [69]
RN [70]
RN [71]
RN [72]
RN [73]
RN [74]
RN [75]
RN [76]
RN [77]
RN [78]
RN [79]
RN [80]
RN [81]
RN [82]
RN [83]
RN [84]
RN [85]
RN [86]
RN [87]
RN [88]
RN [89]
RN [90]
RN [91]
RN [92]
RN [93]
RN [94]
RN [95]
RN [96]
RN [97]
RN [98]
RN [99]
RN [100]
RN [101]
RN [102]
RN [103]
RN [104]
RN [105]
RN [106]
RN [107]
RN [108]
RN [109]
RN [110]
RN [111]
RN [112]
RN [113]
RN [114]
RN [115]
RN [116]
RN [117]
RN [118]
RN [119]
RN [120]
RN [121]
RN [122]
RN [123]
RN [124]
RN [125]
RN [126]
RN [127]
RN [128]
RN [129]
RN [130]
RN [131]
RN [132]
RN [133]
RN [134]
RN [135]
RN [136]
RN [137]
RN [138]
RN [139]
RN [140]
RN [141]
RN [142]
RN [143]
RN [144]
RN [145]
RN [146]
RN [147]
RN [148]
RN [149]
RN [150]
RN [151]
RN [152]
RN [153]
RN [154]
RN [155]
RN [156]
RN [157]
RN [158]
RN [159]
RN [160]
RN [161]
RN [162]
RN [163]
RN [164]
RN [165]
RN [166]
RN [167]
RN [168]
RN [169]
RN [170]
RN [171]
RN [172]
RN [173]
RN [174]
RN [175]
RN [176]
RN [177]
RN [178]
RN [179]
RN [180]
RN [181]
RN [182]
RN [183]
RN [184]
RN [185]
RN [186]
RN [187]
RN [188]
RN [189]
RN [190]
RN [191]
RN [192]
RN [193]
RN [194]
RN [195]
RN [196]
RN [197]
RN [198]
RN [199]
RN [200]
RN [201]
RN [202]
RN [203]
RN [204]
RN [205]
RN [206]
RN [207]
RN [208]
RN [209]
RN [210]
RN [211]
RN [212]
RN [213]
RN [214]
RN [215]
RN [216]
RN [217]
RN [218]
RN [219]
RN [220]
RN [221]
RN [222]
RN [223]
RN [224]
RN [225]
RN [226]
RN [227]
RN [228]
RN [229]
RN [230]
RN [231]
RN [232]
RN [233]
RN [234]
RN [235]
RN [236]
RN [237]
RN [238]
RN [239]
RN [240]
RN [241]
RN [242]
RN [243]
RN [244]
RN [245]
RN [246]
RN [247]
RN [248]
RN [249]
RN [250]
RN [251]
RN [252]
RN [253]
RN [254]
RN [255]
RN [256]
RN [257]
RN [258]
RN [259]
RN [260]
RN [261]
RN [262]
RN [263]
RN [264]
RN [265]
RN [266]
RN [267]
RN [268]
RN [269]
RN [270]
RN [271]
RN [272]
RN [273]
RN [274]
RN [275]
RN [276]
RN [277]
RN [278]
RN [279]
RN [280]
RN [281]
RN [282]
RN [283]
RN [284]
RN [285]
RN [286]
RN [287]
RN [288]
RN [289]
RN [290]
RN [291]
RN [292]
RN [293]
RN [294]
RN [295]
RN [296]
RN [297]
RN [298]
RN [299]
RN [300]
RN [301]
RN [302]
RN [303]
RN [304]
RN [305]
RN [306]
RN [307]
RN [308]
RN [309]
RN [310]
RN [311]
RN [312]
RN [313]
RN [314]
RN [315]
RN [316]
RN [317]
RN [318]
RN [319]
RN [320]
RN [321]
RN [322]
RN [323]
RN [324]
RN [325]
RN [326]
RN [327]
RN [328]
RN [329]
RN [330]
RN [331]
RN [332]
RN [333]
RN [334]
RN [335]
RN [336]
RN [337]
RN [338]
RN [339]
RN [340]
RN [341]
RN [342]
RN [343]
RN [344]
RN [345]
RN [346]
RN [347]
RN [348]
RN [349]
RN [350]
RN [351]
RN [352]
RN [353]
RN [354]
RN [355]
RN [356]
RN [357]
RN [358]
RN [359]
RN [360]
RN [361]
RN [362]
RN [363]
RN [364]
RN [365]
RN [366]
RN [367]
RN [368]
RN [369]
RN [370]
RN [371]
RN [372]
RN [373]
RN [374]
RN [375]
RN [376]
RN [377]
RN [378]
RN [379]
RN [380]
RN [381]
RN [382]
RN [383]
RN [384]
RN [385]
RN [386]
RN [387]
RN [388]
RN [389]
RN [390]
RN [391]
RN [392]
RN [393]
RN [394]
RN [395]
RN [396]
RN [397]
RN [398]
RN [399]
RN [400]
RN [401]
RN [402]
RN [403]
RN [404]
RN [405]
RN [406]
RN [407]
RN [408]
RN [409]
RN [410]
RN [411]
RN [412]
RN [413]
RN [414]
RN [415]
RN [416]
RN [417]
RN [418]
RN [419]
RN [420]
RN [421]
RN [422]
RN [423]
RN [424]
RN [425]
RN [426]
RN [427]
RN [428]
RN [429]
RN [430]
RN [431]
RN [432]
RN [433]
RN [434]
RN [435]
RN [436]
RN [437]
RN [438]
RN [439]
RN [440]
RN [441]
RN [442]
RN [443]
RN [444]
RN [445]
RN [446]
RN [447]
RN [448]
RN [449]
RN [450]
RN [451]
RN [452]
RN [453]
RN [454]
RN [455]
RN [456]
RN [457]
RN [458]
RN [459]
RN [460]
RN [461]
RN [462]
RN [463]
RN [464]
RN [465]
RN [466]
RN [467]
RN [468]
RN [469]
RN [470]
RN [471]
RN [472]
RN [473]
RN [474]
RN [475]
RN [476]
RN [477]
RN [478]
RN [479]
RN [480]
RN [481]
RN [482]
RN [483]
RN [484]
RN [485]
RN [486]
RN [487]
RN [488]
RN [489]
RN [490]
RN [491]
RN [492]
RN [493]
RN [494]
RN [495]
RN [496]
RN [497]
RN [498]
RN [499]
RN [500]
RN [501]
RN [502]
RN [503]
RN [504]
RN [505]
RN [506]
RN [507]
RN [508]
RN [509]
RN [510]
RN [511]
RN [512]
RN [513]
RN [514]
RN [515]
RN [516]
RN [517]
RN [518]
RN [519]
RN [520]
RN [521]
RN [522]
RN [523]
RN [524]
RN [525]
RN [526]
RN [527]
RN [528]
RN [529]
RN [530]
RN [531]
RN [532]
RN [533]
RN [534]
RN [535]
RN [536]
RN [537]
RN [538]
RN [539]
RN [540]
RN [541]
RN [542]
RN [543]
RN [544]
RN [545]
RN [546]
RN [547]
RN [548]
RN [549]
RN [550]
RN [551]
RN [552]
RN [553]
RN [554]
RN [555]
RN [556]
RN [557]
RN [558]
RN [559]
RN [560]
RN [561]
RN [562]
RN [563]
RN [564]
RN [565]
RN [566]
RN [567]
RN [568]
RN [569]
RN [570]
RN [571]
RN [572]
RN [573]
RN [574]
RN [575]
RN [576]
RN [577]
RN [578]
RN [579]
RN [580]
RN [581]
RN [582]
RN [583]
RN [584]
RN [585]
RN [586]
RN [587]
RN [588]
RN [589]
RN [590]
RN [591]
RN [592]
RN [593]
RN [594]
RN [595]
RN [596]
RN [597]
RN [598]
RN [599]
RN [600]
RN [601]
RN [602]
RN [603]
RN [604]
RN [605]
RN [606]
RN [607]
RN [608]
RN [609]
RN [610]
RN [611]
RN [612]
RN [613]
RN [614]
RN [615]
RN [616]
RN [617]
RN [618]
RN [619]
RN [620]
RN [621]
RN [622]
RN [623]
RN [624]
RN [625]
RN [626]
RN [627]
RN [628]
RN [629]
RN [630]
RN [631]
RN [632]
RN [633]
RN [634]
RN [635]
RN [636]
RN [637]
RN [638]
RN [639]
RN [640]
RN [641]
RN [642]
RN [643]
RN [644]
RN [645]
RN [646]
RN [647]
RN [648]
RN [649]
RN [650]
RN [651]
RN [652]
RN [653]
RN [654]
RN [655]
RN [656]
RN [657]
RN [658]
RN [659]
RN [660]
RN [661]
RN [662]
RN [663]
RN [664]
RN [665]
RN [666]
RN [667]
RN [668]
RN [669]
RN [670]
RN [671]
RN [672]
RN [673]
RN [674]
RN [675]
RN [676]
RN [677]
RN [678]
RN [679]
RN [680]
RN [681]
RN [682]
RN [683]
RN [684]
RN [685]
RN [686]
RN [687]
RN [688]
RN [689]
RN [690]
RN [691]
RN [692]
RN [693]
RN [694]
RN [695]
RN [696]
RN [697]
RN [698]
RN [699]
RN [700]
RN [701]
RN [702]
RN [703]
RN [704]
RN [705]
RN [706]
RN [707]
RN [708]
RN [709]
RN [710]
RN [711]
RN [712]
RN [713]
RN [714]
RN [715]
RN [716]
RN [717]
RN [718]
RN [719]
RN [720]
RN [721]
RN [722]
RN [723]
RN [724]
RN [725]
RN [726]
RN [727]
RN [728]
RN [729]
RN [730]
RN [731]
RN [732]
RN [733]
RN [734]
RN [735]
RN [736]
RN [737]
RN [738]
RN [739]
RN [740]
RN [741]
RN [742]
RN [743]
RN [744]
RN [745]
RN [746]
RN [747]
RN [748]
RN [749]
RN [750]
RN [751]
RN [752]
RN [753]
RN [754]
RN [755]
RN [756]
RN [757]
RN [758]
RN [759]
RN [760]
RN [761]
RN [762]
RN [763]
RN [764]
RN [765]
RN [766]
RN [767]
RN [768]
RN [769]
RN [770]
RN [771]
RN [772]
RN [773]
RN [774]
RN [775]
RN [776]
RN [777]
RN [778]
RN [779]
RN [780]
RN [781]
RN [782]
RN [783]
RN [784]
RN [785]
RN [786]
RN [787]
RN [788]
RN [789]
RN [790]
RN [791]
RN [792]
RN [793]
RN [794]
RN [795]
RN [796]
RN [797]
RN [798]
RN [799]
RN [800]
RN [801]
RN [802]
RN [803]
RN [804]
RN [805]
RN [806]
RN [807]
RN [808]
RN [809]
RN [810]
RN [811]
RN [812]
RN [813]
RN [814]
RN [815]
RN [816]
RN [817]
RN [818]
RN [819]
RN [820]
RN [821]
RN [822]
RN [823]
RN [824]
RN [825]
RN [826]
RN [827]
RN [828]
RN [829]
RN [830]
RN [831]
RN [832]
RN [833]
RN [834]
RN [835]
RN [836]
RN [837]
RN [838]
RN [839]
RN [840]
RN [841]
RN [842]
RN [843]
RN [844]
RN [845]
RN [846]
RN [847]
RN [848]
RN [849]
RN [850]
RN [851]
RN [852]
RN [853]
RN [854]
RN [855]
RN [856]
RN [857]
RN [858]
RN [859]
RN [860]
RN [861]
RN [862]
RN [863]
RN [864]
RN [865]
RN [866]
RN [867]
RN [868]
RN [869]
RN [870]
RN [871]
RN [872]
RN [873]
RN [874]
RN [875]
RN [876]
RN [877]
RN [878]
RN [879]
RN [880]
RN [881]
RN [882]
RN [883]
RN [884]
RN [885]
RN [886]
RN [887]
RN [888]
RN [889]
RN [890]
RN [891]
RN [892]
RN [893]
RN [894]
RN [895]
RN [896]
RN [897]
RN [898]
RN [899]
RN [900]
RN [901]
RN [902]
RN [903]
RN [904]
RN [905]
RN [906]
RN [907]
RN [908]
RN [909]
RN [910]
RN [911]
RN [912]
RN [913]
RN [914]
RN [915]
RN [916]
RN [917]
RN [918]
RN [919]
RN [920]
RN [921]
RN [922]
RN [923]
RN [924]
RN [925]
RN [926]
RN [927]
RN [928]
RN [929]
RN [930]
RN [931]
RN [932]
RN [933]
RN [934]
RN [935]
RN [936]
RN [937]
RN [938]
RN [939]
RN [940]
RN [941]
RN [942]
RN [943]
RN [944]
RN [945]
RN [946]
RN [947]
RN [948]
RN [949]
RN [950]
RN [951]
RN [952]
RN [953]
RN [954]
RN [955]
RN [956]
RN [957]
RN [958]
RN [959]
RN [960]
RN [961]
RN [962]
RN [963]
RN [964]
RN [965]
RN [966]
RN [967]
RN [968]
RN [969]
RN [970]
RN [971]
RN [972]
RN [973]
RN [974]
RN [975]
RN [976]
RN [977]
RN [978]
RN [979]
RN [980]
RN [981]
RN [982]
RN [983]
RN [984]
RN [985]
RN [986]
RN [987]
RN [988]
RN [989]
RN [990]
RN [991]
RN [992]
RN [993]
RN [994]
RN [995]
RN [996]
RN [997]
RN [998]
RN [999]
RN [1000]

```

DR Pfam; PF00560; LRR; 10.
DR Pfam; PF01463; LRRCT; 1.
DR Pfam; PF01462; LRRNT; 1.
DR PRINTS; PR00019; LEURICHRPT.
DR SMART; SM00060; FN3; 1.
DR SMART; SM00408; IGC2; 1.
DR SMART; SM00370; LRR; 4.
DR SMART; SM00082; LRRCT; 1.
DR SMART; SM00013; LRRNT; 1.
DR SMART; SM00369; LRR_TYP; 1.
KW Immunoglobulin domain; Signal.
FT SIGNAL 1 20
FT CHAIN 21 >707
FT NON_TER 707 707
SQ SEQUENCE 707 AA; 79156 MW; F4E7C1573DD165B CRC64;

Query Match 11.6%; Score 394.5; DB 11; Length 707;
Best Local Similarity 22.6%; Pred. No. 1.3e-20;
Matches 166; Conservative 116; Mismatches 258; Indels 193; Gaps 31;

QY 10 LIGKIGLQVAPLSVMAKSCPSVCR-----DAGFYCNDRTFTSPTGI 56
DB 11 LLGLATITLVAQDKV--DCPQLCTCEIRPWFPTPSYMEASTVDCNDLGLINPAKL 67

QY 57 PEDATLYLQNNQI-----NNAGIPSLKMLKVERI-----YLYHNS 94
DB 68 PADTQILLQTNRIAREHSTDFPVLTLGLDLSQNNLSSTVNVNQMQLSLLVYLEENK 127

QY 95 LDEP-----TNLPK-YVK-----ELHLOENNIRTIYDSLSKI 127
DB 128 LTELPEKCLYGSNQLVYVNNHLLSTSPGAFIGLHNLRLHLSNKLQNSQWFDAL 187

QY 128 PYLEELHDDNSAVSIEGAFRDSNYLRLLEFSRNLSTIP-----WGLPRTIEELRD 183
DB 188 PNLEILMGDNP1--IRIKMNFQPKLSRSLVIAGINTEIPDLAGL-ENLESISY 244

QY 184 DNRISTSSPSLOGLTSKRLVLDGNNHGLGDKVFNVLNVLTELSVNSI-----237
DB 245 DNRLSKYPOVALQKAVNLKFLDKNPNIRIRGD--FSNMLHLKELGI--NNMPLVSI 300

QY 238 -TAAPVNLPTNLRKLYLQDN-HNRYPPNAPSYLRQLYRLDMSNNLSNLPQIGFDLD 295
DB 301 DSLAVDMLP--DLRKIEATNPRISYTHNAPFELPKLESIMLVNALSALYHGTIESLP 358

QY 296 NITOLILRNPNWYCGCKMKVRDWLQSLPKVNVN-----GLMCOAPEKVGMAIKDLN- 349
DB 359 NLKEISTHNPICDCVIRWIN-----MNKTNIREFMPSLFCVDPPEFQGNVRQVHR 412

QY 350 -----AELF-----DCKDSGVSTIQTITA-----IPNT-----373
DB 413 RDMMEICLPPIAPESFSDLDVEADSVSLHCRATAEPOPEYIWTTPGKKLPNTMREK 472

QY 374 -----VTPAQGNFAPVTKQ---PDKNPKL-----TKDQQTGSPSRKTIIT 413
DB 473 FYVHSEGVLSIRGITPEKGGYTCIATNLVGCADLSIMIKVGGSVPDNNGS-----LNI 527

QY 414 TVKSVTSDTHISWKLAPMTALBSLWKLGHSPAGSGTITIVTGERSELYVTALEPDS 473
DB 528 KIRDIRANSVLVSWKASSILKSKYKVTGFTKEDSAAQASARIPIVSDYKVTNLTHLKPST 587

QY 474 PYKVMVPMETSNLYLDFETPVCITETAPLRMKNPTTLNREOKPEKPKNPLPLAAII 533
DB 588 EYKIC---IDIPVYQ-KSRQCVN-----VTKSLHOGKYEYKXNHVTFVACV 632

QY 534 GAVALVTIALLALVCWY--VHRNGSLFSRNCAYSKGRKKDDYAEA-----GT 580
DB 633 GGLLGIY-----GVMCLFSCVQSGSGEGH-SYAVNHCHKPALAFSELYPPLINLWESS 686

QY 581 KKDNSILEIRETS 593
DB 687 KEKRATLEVAKATA 699

RESULT 9
Q9HCJ2 PRELIMINARY; PRT; 640 AA.
ID Q9HCJ2
AC Q9HCJ2;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE KIAA1580 protein (Fragment).
GN KIAA1580.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20450863; PubMed=10997877;
RA Nagase T., Kikuno R., Nakayama M., Hirose M., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes.
RT XVIII. The complete sequences of 100 new cDNA clones from brain which
RT code for large proteins in vitro.";
RL DNA Res. 7:273-281(2000).
DR EXEL; A5046800; BAB13406.1; -.
DR HSP; P22888; ILUT.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003598; Ig_c2.
DR InterPro; IPR003600; Ig_like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR000483; LRR_Cterm.
DR InterPro; IPR000372; LRR_Nterm.
DR InterPro; IPR003592; LRR_out.
DR InterPro; IPR003591; LRR_Typ.
DR Pfam; PF00047; Ig; 1.
DR Pfam; PF00560; LRR; 9.
DR Pfam; PF01463; LRRCT; 1.
DR Pfam; PF01462; LRRNT; 1.
DR PRINTS; PR00019; LEURICHRPT.
DR SMART; SM00408; IGC2; 1.
DR SMART; SM00410; IG_Like; 1.
DR SMART; SM00370; LRR; 6.
DR SMART; SM00082; LRRCT; 1.
DR SMART; SM00013; LRRNT; 1.
DR SMART; SM00369; LRR_TYP; 8.
KW Immunoglobulin domain.
FT NON_TER 1
SQ SEQUENCE 640 AA; 71949 MW; 6058974872636838 CRC64;

Query Match 11.3%; Score 385; DB 4; Length 640;
Best Local Similarity 23.8%; Pred. No. 5.6e-20;
Matches 163; Conservative 98; Mismatches 256; Indels 169; Gaps 23;

QY 10 LIGKIGLQVAPLSVMAKSCPSVCRDAGF--IYCNDRFTSPTGIPEDATTLYLQN 67
DB 26 LLVLLALQLLVVAGVRAQTCPSVCSNQFSKYICVRKNLREVDPGISTNTLNLHE 85

QY 68 NQINNAGIPSDKLNKLKVERIYLYHNSLDEFPNLPKYVKELHLOENNIRTIYDSLSKI 127
DB 86 NQI-----QIYK-----NSFKHL-----RHELIQLSRNHIRTETIGAFNGL 123

QY 128 PYLEELHDDNSAVSIEGAFRDSNYLRLLEFSRNLSTIPWGLPRTIEELRD--- 184
DB 124 ANLMTLELFDNLT--TIPNGAFVYLSKLKELWLRNPNPESIPSYAFNRPISLRRDLGE 181

QY 185 -NRISTSSPSLOGLTSKRLVLD--GNLNNHGLGDKVFFVNLVNLTELSILVNSLTA-AP 241
DB 182 LKRSLYSRGAPEGLSNLYNLNLAACNREIPNTP-----LIKDELDSLGNLSAIRP 236

QY 242 VNLPG-TNLRLKLYLDQNNHINRVPPNAPSYLRQLYRLDMSNNLSNLPQIGFDLDNITOL 300
DB 237 GSFOGLMHLQKLMQISQIQVIERNAFDNLQSLIVEINLAHNNLILPDLFTPLHLERI 296

QY 301 ILRNPNWYCGCKMKVVRDWLQSLPKVNVNVEGLMCQAPKVRGMAIKDLNLELFDCKDSGI 360

J

Db 190 NNNISRLVTSFNMPKIRTLRHSNHLVCDCHLAWLSWLRQRRRTIGQFTLCMAPVHLR 249
 QY 166 -----LSIPWGLPRTI 177
 Db 250 GFSVADYQKKEYVCPGPHSEAPACNANSLSCPSACSNIVDCRGKGLTEIPANLPEGI 309
 QY 178 BELRLDDNRISTSSPSLQGLSLKRLVLDGNLNNHGLGDKVFFNLVNLVLSLVRNSL 237
 Db 310 VEIRLEQNSIKSPAGAFQYKKLRIDISKNOISD--IAPDAFQGLKSLTSLVLYGNKI 367
 QY 238 TAAPVNL-PG-TNLRKLYLQDNHINRPPNFAVSYLRLYELDMNNNLNLPQGLFDDLD 295
 Db 368 TEIPKGLFDGLVSLQLLNANKINCLRVNTFQDLQNLNLSLYDNKLOTISKGLFAPLQ 427
 QY 296 NITOLILNPNWYCGCKMKWVROWLQSLPVKVNVRGLMCOAPEKVRGMAIKDLNAELFDC 355
 Db 428 SIOTLHLAQNPFVCDCHLAWLADYLDNPIETS--GARCSPRRLANKRISQIKSKKFC 485
 QY 356 KDSGIVSTIQTITTAIPNTVYPAGOWPAPVTKQPDINKPLTKDQQTGSPSKRTITIV 415
 Db 486 SGSEYRNRFSSFCFMDLVCPEKCEGTIV---DCSNQKLAR-----IPSHLPEYTTD 536
 QY 416 KSVTSDTHISWKLALPMTAL-----ELSWKLGHSPAPGSGTETIVTGE 460
 Db 537 LRINDNDISV-----LEATGIFKKLPNLRKINLSNRIKEVREGADGAASVQELMLTGN 591
 QY 461 RSEYLVITAL-EPDSPYKVMVPMETSNLYLDFETPVCIETETAP-----LRMY-NEPTT 511
 Db 592 QLETHMGFRGLSGLKTLML--RSNLI-----SCVSNDFAGLSVRLISLYD--- 638
 QY 513 INREGEKPKNPNLPLAIIIGA---VALVTIALLA---LVVYVHRNGSLFSRNCA 564
 Db 639 -NR-----ITHTTGAFTVLSLSTINLSNPNFCNCHMAWLGRLKRIIV 684
 QY 565 YSKGRRRRDYAEAGTKDKNSILEI-----RETFOMLPISNPEIKSEEVIL-----H 612
 Db 685 SONPCQKPFELKEIPIQDVATQDFTCDGNEESSQLSPRCPEQFTCVETVVRCSNRLH 744
 QY 613 TIFPNMG-----NLY-KNNH 627
 Db 745 AL--PKGMPKDVTELYEGNH 763
 RESULT 13
 ID 088280 PRELIMINARY; PRG; 1523 AA.
 AC 088280;
 DT 01-NOV-1998 (TrEMBLrel. 08, Created)
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE MEGF5.
 GN MEGF5.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SPRAGUE-DAWLEY; TISSUE=BRAIN;
 RX MEDLINE=98360089; PubMed=9693030;
 RA Nakayama M., Nakajima D., Nagase T., Nomura N., Seki N., Ohara O.:
 RT "Identification of high-molecular-weight proteins with multiple EGF-
 RL like motifs by motif-crap screening."
 RL Genomics 51:27-34(1998).
 DR DR
 DR HSP; P01132; 1EGF.
 DR InterPro; IPR000152; ASX_hydroxyl.
 DR InterPro; IPR000359; Cys_knot.
 DR InterPro; IPR000561; EGF_like.
 DR InterPro; IPR000742; EGF_2.
 DR InterPro; IPR001881; EGF_Ca.
 DR InterPro; IPR002400; GF_cysknot.
 DR InterPro; IPR001791; Laminin_G.

DR InterPro; IPR001611; LRR.
 DR InterPro; IPR000483; LRR_Cterm.
 DR InterPro; IPR000372; LRR_Nterm.
 DR InterPro; IPR003592; LRR_out.
 DR InterPro; IPR003591; LRR_Typ.
 DR Pfam; PF00008; EGF; 9.
 DR Pfam; PF00054; laminin_G; 1.
 DR Pfam; PF00056; LRR; 19.
 DR Pfam; PF01463; LRRCT; 4.
 DR Pfam; PF01462; LRRNT; 4.
 DR PRINTS; PR00438; GFCYSKNOT.
 DR SMART; SM00441; CT; 1.
 DR SMART; SM00179; EGF_Ca; 1.
 DR SMART; SM00001; EGF_like; 8.
 DR SMART; SM00282; Lamg; 1.
 DR SMART; SM00370; LRR; 5.
 DR SMART; SM00082; LRRCT; 4.
 DR SMART; SM00013; LRRNT; 4.
 DR SMART; SM00369; LRR_Typ; 9.
 DR PROSITE; PS00010; ASX_HYDROXYL; UNKNOWN_2.
 DR PROSITE; PS01185; CTCK_1; UNKNOWN_1.
 DR PROSITE; PS01225; CTCK_2; 1.
 DR PROSITE; PS00022; EGF_1; UNKNOWN_9.
 DR PROSITE; PS01186; EGF_2; 7.
 DR PROSITE; PS01187; EGF_Ca; 2.
 KW Calcium-binding; EGF-like domain; Glycoprotein; Repeat.
 SQ SEQUENCE 1523 AA; 167767 MW; 6CE1B7AF9244478E CRC64;
 Query Match 11.2%; Score 381.5; DB 11; Length 1523;
 Best Local Similarity 22.3%; Pred. No. 3.5e-19;
 Matches 179; Conservative 107; Mismatches 272; Indels 243; Gaps 29;
 QY 17 LFLQVAPVLSV-----AKSCPSVCRCDAGFIYCNDRFLTSTIPTGIPEDA----- 60
 Db 16 LALALALASILSPAPAACTCTCSAASVDCHGUGLRAVPRGIPRNAEELDRNNITR 75
 QY 61 -----TTLVYQNNQIN--NAGIPSDLANLKVRIYLYHNSLDEPPNL----- 102
 Db 76 ITKMDFTGLKRLVHLEDNQSVIERGAFODLKQL---EELRLNKNKIQVPELLFQST 132
 QY 103 PKYVRELHLOENRITITVDSLKIPYLEELHDDNSVSAVSTEEGAFR----- 151
 Db 133 PKLTR-LDLSNQIQGIPKAFRGVTVGNLQDNHHSIC--IEDGAFRALRDLLEILTLN 189
 QY 152 DSNYLRLLFLSRNH----- 165
 Db 190 NNNISRLVTSFNMPKIRTLRHSNHLVCDCHLAWLSWLRQRRRTIGQFTLCMAPVHLR 249
 QY 166 -----LSIPWGLPRTI 177
 Db 250 GFSVADYQKKEYVCPGPHSEAPACNANSLSCPSACSNIVDCRGKGLTEIPANLPEGI 309
 QY 178 BELRLDDNRISTSSPSLQGLSLKRLVLDGNLNNHGLGDKVFFNLVNLVLSLVRNSL 237
 Db 310 VEIRLEQNSIKSPAGAFQYKKLRIDISKNOISD--IAPDAFQGLKSLTSLVLYGNKI 367
 QY 238 TAAPVNL-PG-TNLRKLYLQDNHINRPPNFAVSYLRLYELDMNNNLNLPQGLFDDLD 295
 Db 368 TEIPKGLFDGLVSLQLLNANKINCLRVNTFQDLQNLNLSLYDNKLOTISKGLFAPLQ 427
 QY 296 NITOLILNPNWYCGCKMKWVROWLQSLPVKVNVRGLMCOAPEKVRGMAIKDLNAELFDC 355
 Db 428 SIOTLHLAQNPFVCDCHLAWLADYLDNPIETS--GARCSPRRLANKRISQIKSKKFC 485
 QY 356 KDSGIVSTIQTITTAIPNTVYPAGOWPAPVTKQPDINKPLTKDQQTGSPSKRTITIV 415
 Db 486 SGSEYRNRFSSFCFMDLVCPEKCEGTIV---DCSNQKLAR-----IPSHLPEYTTD 536
 QY 416 KSVTSDTHISWKLALPMTAL-----ELSWKLGHSPAPGSGTETIVTGE 460
 Db 537 LRINDNDISV-----LEATGIFKKLPNLRKINLSNRIKEVREGADGAASVQELMLTGN 591
 QY 461 RSEYLVITAL-EPDSPYKVMVPMETSNLYLDFETPVCIETETAP-----LRMY-NEPTT 511


```

DR SMART; SM00409; IG; 1.
DR SMART; SM00408; IGC2; 1.
DR SMART; SM00410; IG like; 1.
DR SMART; SM00370; LRR; 7.
DR SMART; SM00082; LRRCT; 1.
DR SMART; SM00013; LRRNT; 1.
DR SMART; SM00369; LRR_TYP; 8.
KW Immunoglobulin domain.
SQ SEQUENCE 707 AA; 79064 MW; 26A210F671EDA875 CRC64;

Query Match      11.1%; Score 376.5; DB 11; Length 707;
Best Local Similarity 22.0%; Pred. No. 2.7e-19;
Matches 162; Conservative 118; Mismatches 256; Indels 201; Gaps 31;

QY 10 LIGTKIGLFLQVAPLSVMAKSPSCVRC-----DAGFYCNDRELSIPTGI 56
DB 11 LIGTAIVAGAGKKV---DQPLCTCEIRPWTTPRSIYWEASTVDCNDLGLLNFARL 67
QY 57 PEDATLYLQNNQI-----NNAGIPSDLNKLLKYRI-----YLHNS 94
DB 68 PADTQILLLOTNNIARIETHSDFPVNLAGLDLSONLSSVTNINQKMSQLSVYLEENK 127
QY 95 LDEEP-----TNLPK-VYK-----ELHLOENNRITTYDSLKI 127
DB 128 LTELPEKCLGLSLQLGLVYNNLLSAPGAFVGLHNLRLHLSNRLOMINSKWFEL 187
QY 128 PYLEELHDDNSVANSIEEGAFRDSNYRLIFLSRNHLSTIP---WGLPTTIEELKLD 183
DB 188 PNLEITMLGDNPI--LRKDMNEPQLLKLSRVIAGINLVEPDDALVGL-ENLESISFY 244
QY 184 DNRISTISSPSLQGLSLKRLVDGNLNNHGLGDKVFFNLVNLTELSVRNSL----- 237
DB 245 DNRILKVPQALQKAVNLKFLDKNPNIRIRGD--FSNMLHKLGI--NNMPELWSI 300
QY 238 -TAAEPVNLPGNFKLYLODN-HINRVPNPAFSLYRLQYRLDMSNNLSNLPQGIFFDLD 295
DB 301 DSLAVDNLP--DLRKIRATNNPLSLIHPNAFFRLPKLESIMLSNLSALYHGTIESLP 358
QY 296 NITQILIRNNFWYCGCKMKWVRDLQSLFVKVYVR-----GLMCAPEKVRGMAIKDLN- 349
DB 359 NLKEISIHNFIRDCVIRWIN-----MNKTINIFEMEPDLSFCVDPPEFGQNVQVHF 412
QY 350 -----AELEDCXDSIVSTIQITTA-----IPNT----- 373
DB 413 RDMMEICLPLIAPESFSLDVEADSVSLHCRATAEPQEIYWIWTPSGKRLLPNTLREK 472
QY 374 -----VYPAAGOWPAPYTKQ---PDINKPLA---KDQTTGSPSRKTTI 413
DB 473 FYVHSEGTLDIRGTPYKGGLYTCIATNLVGADLKSINIKVGFVPQDNNGS-----LNI 527
QY 414 TYKSVTSDTHISWKLALPMTALRLSWLKLGHSPAFGSITETIVTGERSEYLVTALEPDS 473
DB 528 KIRDIRANSVLVSWKANSKILSKSVKWTAFVKTEDSQAAQSARIPSDVKVYNLTHLKPST 587
QY 474 PYKVCWPMETSNLYLDETPVCIETETAPLRYNPTTLNREQEKPEPKENLPLAII 533
DB 588 EYKIC---IDITPIY-----QSKRKQCVNVVTKSLEHDGKNGKSHIV-FVACY 632
QY 534 GGAVALVTIALLALVCWY--VHRNGSLFSRNC---AVSKGRRRKDDYAEA----- 578
DB 633 GG-----LLGIIGVWCLFGCVSQSEG-----NCENEHSYTVNHCHKPTLAFSELYPPLNL 682
QY 579 --GTKDMSIILEIRTS 593
DB 683 WESSKEKPALEVKATA 699

```

Search completed: April 23, 2003, 11:06:21
Job time : 41 secs